IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF MASSACHUSETTS

MASSACHUSETTS INSTITUTE OF TECHNOLOGY,

Plaintiff,

v.

HARMAN INTERNATIONAL INDUSTRIES, INCORPORATED,

Defendant.

Civil Action No. 05-10990-DPW
PUBLIC VERSION

HARMAN'S MEMORANDUM IN SUPPORT OF ITS MOTION FOR SUMMARY JUDGMENT THAT CLAIMS 1, 42 AND 45 OF THE '685 PATENT ARE INVALID UNDER 35 U.S.C. § 102(b) DUE TO PUBLIC USE

TABLE OF CONTENTS

00171503.DOC /

TABLE OF AUTHORITIES

Page(s)

Federal Cases

Page(s)

Harman is entitled to summary judgment of invalidity under Section 102(b) because there is no genuine issue of material fact that from May to July of 1989, at least fifty persons (including students and third-party corporate representatives from various corporations) used MIT's Back Seat Driver system to drive around the Boston area at least fifty times. Each use included the subject matters of claims 1, 42, and 45 of U.S. Patent No. 5,177,685 ("the '685 patent"). Each of these fifty uses occurred "more than one year prior to the date of [MIT's] application for patent in the United States."[1] 35 U.S.C. § 102(b). None of those uses was subject to any obligation of confidentiality.[2] And, as a matter of law, none of those uses was experimental. Consequently, each of them constitutes a "public use" in violation of Section 102(b) of the Patent Act. 35 U.S.C. § 102(b) ("A person shall be entitled to a patent unless . . . (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of the application for patent in the United States").

Section 102(b) prevents monopolization of knowledge that is in public use. 35 U.S.C. § 102(b); *Bonito Boats, Inc. v. Thunder Craft Boats, Inc.*, 489 U.S. 141, 148 (1989). To prevail on summary judgment of invalidity of claims 1, 42, and 45 of the '685 patent Harman need only prove that *one* of those fifty public uses occurred. *Egbert v. Lippmann*, 104 U.S. 333,

---

1 That date which is exactly "one year prior to the date of the application for patent" is referred to as the "critical date."

2 Indeed, MIT does not dispute, in its Opposition To Harman's Motion For Summary Judgment Of Unenforceability Of The '685 Patent, that: (1) students and corporate representatives used the working Back Seat Driver system at least fifty times on the streets of Boston; (2) prior to the critical date; and (3) no confidentiality agreements were signed regarding any of those uses. (*See* MIT's Counter-Statement Of Facts Pursuant To Local Rule 56.1 In Opposition To Harman's Motion For Summary Judgment Of Unenforceability Of The '685 Patent, Docket Entry No. 146, at pp. 16 - 17).

336 (1881) ("one well-defined case of such use is just as effectual to annul the patent as many"). The undisputed evidence presented here meets this burden.

Here, MIT's own documents, the named inventors' sworn deposition testimony, MIT's sworn 30(b)(6) testimony, and MIT's written interrogatory responses prove that prior to the statutory "critical date" MIT performed ***fifty public uses*** of the subject matter of claims 1, 42, and 45, as shown below:

Date	**Event**	**Evidence**
At least as early as June 1989	Claim 1 reduced to practice	SOF 5, 10, 14-17
At least as early as June 1989	Claim 42 reduced to practice	SOF 18-24, 28-31
At least as early as June 1989	Claim 45 reduced to practice	SOF 19-20, 34-39
From May 1, 1989 to July 31, 1989	Fifty public uses of the Back Seat Driver by fifty different people on the public streets of Boston, including the subject matter of claims 1, 42, and 45	SOF 6, 10-14, 19-24, 34-35
August 9, 1989	"Critical date" for § 102(b)	SOF 1
August 9, 1990	Application for '685 patent filed	SOF 1

Each of these uses involved third parties. Yet, none of these public uses was confidential. *See* Section IA. Instead, MIT freely and widely publicized its Back Seat Driver system, including these pre-critical-date public uses. *See* Section IA. In addition, at least some of MIT's public uses were for commercial purposes. *See* Section IB. These facts demonstrate that each of these uses was public. Importantly, none of these uses was "experimental" because claims 1, 42, and 45 were already reduced to practice prior to each of these public uses. (SOF 15, 28-31, 36-38.) *See* Section III; *see also, Allen Eng'g Corp. v. Bartell Indus., Inc.*, 299 F.3d 1336, 1354

(Fed. Cir. 2002) (holding "once the invention is reduced to practice, there can be no experimental use negation.") Accordingly, the sole exception to the statutory bar of § 102(b) is unavailable to MIT, as a matter of law.

Because there is no genuine issue of material fact that at least one of MIT's fifty exhibitions was a public use in violation of Section 102(b), Harman respectfully submits that summary judgment of invalidity of claims 1, 42, and 45 of the '685 patent is warranted.

SUMMARY OF APPLICABLE LAW

Invalidity is a question of law amenable to summary judgment. *Baxter Int'l., Inc. v. COBE Labs., Inc.*, 88 F.3d 1054, 1058 (Fed. Cir. 1996) (affirming summary judgment of invalidity under Section 102(b) due to public use); *accord Netscape Commc'ns Corp. v. Konrad*, 295 F.3d 1315, 1324-25 (Fed. Cir. 2002); *New Railhead Mfg. v. Vermeer Mfg. Co.*, 298 F.3d 1290 (Fed. Cir. 2002); *Lockwood v. American Airlines, Inc.*, 107 F.3d 1565 (Fed. Cir. 1997).

Summary judgment is appropriate here because there is no genuine issue of material fact, and Harman is entitled to judgment as a matter of law. *See* FED. R. CIV. P. 56(c); *Lockwood*, 107 F.3d at 1569. "The mere existence of a scintilla of evidence in support of [MIT's] position will be insufficient; there must be evidence on which the jury could reasonably find for [MIT]." *Anderson v. Liberty Lobby, Inc.*, 477 U.S. 242, 252 (1986). MIT must set forth specific facts showing that there is a genuine issue for trial. FED. R. CIV. P. 56(e). There is no such evidence in the record, and thus Harman is entitled to judgment in its favor. This Court's application of the statute is thus simple, straight-forward and involves no issue of claim construction.[3] MIT

3 Claim 1 presents no issue of claim construction because MIT concedes that the subject matter of claim 1 was present and reduced to practice in the relevant, pre-critical-date uses. *See* Section IIA. The parties do not dispute any aspect of the construction of claim 42. With respect to claim 45, for purposes of this motion Harman takes MIT at their word that claim 45 is properly construed to cover second, shorter instructions that

publicly used its Back Seat Driver, including the subject matters of claims 1, 42, and 45, after they were reduced to practice and before August 9, 1989, but did not apply for a patent until more than a year later, on August 9, 1990. Accordingly, claims 1, 42, and 45 of MIT's patent are invalid.

ARGUMENT

I. Fifty Public Uses Of MIT's Back Seat Driver Occurred In The United States More Than One Year Before MIT Filed Its Patent Application.

By April 30, 1989, ([[**REDACTED**]])

(Tab 4, filed under seal, SOF 4)

By June, 1989, MIT's Back Seat Driver was "a working system on the road," (SOF 6) and had already been "successfully used by drivers who ha[d] never driven in Boston." (SOF 10) Between May 1 and July 31, 1989, at least fifty people used MIT's Back Seat Driver to successfully navigate around the public streets of Boston. (SOF 6) The people known to have used and witnessed the use of MIT's Back Seat Driver include a Bellcore employee, NEC employees, several MIT undergraduate students, and several General Motors employees. (SOF 6, 7, 53)

As a matter of law, these uses on the public streets of Boston were "public" in the sense of § 102(b). "Public use" under § 102(b) does not require use that is open and visible to the public. *New Railhead*, 298 F.3d at 1297 (citing *Lough v. Brunswick Corp.*, 86 F.3d 1113, 1119 (Fed. Cir. 1996)). Nor does "public use" require a use that discloses the invention to the public. *In re Epstein*, 32 F.3d 1559, 1568 (Fed. Cir. 1994). Rather "public use" under § 102(b) simply means that the invention was used in public "in its natural [or] intended way," nothing more.

---

are merely "near" (rather than "at") the place to act, and the underlying contemporaneous documents demonstrate that such subject matter was present and reduced to practice in the relevant, pre-critical-date uses. *See* Section IIC.

Sys. Mgmt. Arts Inc. v. Avesta Techs., Inc., 87 F. Supp. 2d 258, 270 (S.D.N.Y. 2000); *see also Hall v. Macneale*, 107 U.S. 90 (1883); *Egbert*, 104 U.S. at 336. This has been "black letter law" since the famous "Corset Case" in 1881, in which the United States Supreme Court found that a "public use" barring patentability occurs even when the claimed invention is concealed from public view. *See Egbert*, 104 U.S. at 336 (a corset worn in public was in "public use," thereby invalidating the patent, even though the invention was concealed within the corset); *see also Hall*, 107 U.S. at 97 (a safe mechanism was in public use even though the invention could not be seen without destroying the safe); *Netscape*, 295 F.3d at 1319 (affirming summary judgment of "public use" invalidity in light of a single demonstration to two people within a university computing lab, without any obligation of confidentiality); *New Railhead*, 298 F. 3d at 1298 (affirming summary judgment of "public use" invalidity even though the use was by a single person, was underground beneath public land next to an interstate highway, was out of the view of the public, was secret and confidential, and could not be seen by the public, in general).[4]

As a matter of law, MIT's fifty uses of the Back Seat Driver on the public streets of Boston were public. (SOF 6) It makes no difference under § 102(b) whether any pedestrians or other drivers saw or heard those fifty uses. All that matters is that at least one of the fifty different drivers used the system in public, in an automobile, on city streets, as the alleged invention was intended to be used. *See In re Epstein*, 32 F.3d at 1568; *see also*, *New Railhead*, 298 F.3d at 1299 ("[i]t is not public knowledge of his invention that precludes the inventor from obtaining a patent for it, but a public use . . . .").[5]

---

4 *See also Lough*, 86 F.3d at 1119; *Epstein*, 32 F.3d at 1568; *Sys. Mgmt. Arts*, 87 F. Supp. 2d at 268; *Minn. Mining and Mfg. Co. v. Appleton Papers, Inc.*, 35 F. Supp. 2d 1138, 1148 (D. Minn. 1999) ("*3M*").

5 Nor does it matter that the working system relied on parts that resided outside of the vehicle in the MIT Media Lab. *See NTP, Inc. v. Research In Motion*, Ltd., 418 F.3d 1282, 1317 (Fed. Cir. 2005) ("The use of a claimed system under section 271(a) is the place at which the system as a whole is put into service, i.e., the place where control of the system is exercised and beneficial use of the system obtained.") Whether "elements

It is likewise irrelevant under § 102(b) whether any of the fifty users understood how the system worked as a result of their fifty uses. There is no requirement under § 102(b) for an enablement-type inquiry into whether the public learned about or understood the technology as a result of the use. *Netscape*, 295 F.3d at 1323. The relevant question is whether the invention was *used in public*, ***not*** whether the use disclosed the invention to the public, enabled the public to comprehend or understand the invention, or even whether the public could see the invention. *Epstein*, 32 F.3d at 1568.[6]

A. None of The Pre-Critical-Date Public Uses of MIT's Back Seat Driver Was Confidential.

There is no genuine dispute that MIT neither obtained nor enforced any confidentiality obligations in connection with the fifty pre-critical-date uses of its Back Seat Driver system. (SOF 8) MIT's failure to keep the system confidential further demonstrates the invalidity of claims 1, 42, and 45 under § 102(b). (SOF 8, 9, 43, 50-51, 56) The absence of confidentiality is strongly indicative of public use. In *Baxter Int'l, Inc. v. COBE Labs., Inc.*, for example, the Federal Circuit affirmed summary judgment of invalidity under the "public use" bar of § 102(b) and rejected the patentee's argument that confidentiality should be implied where the use occurred under ethical obligations to refrain from taking credit for the work of others, or from publishing the work of others without permission. *See Baxter*, 88 F.3d 1059 (Fed. Cir. 1996) (finding public use due, in part, to policy of allowing "free flow" of information). Under *Baxter*,

---

of the 'invention' were *inside* the car and entirely out of view of the public" is entirely beside the point. The Back Seat Driver was used in public "in its natural or intended way," and such use is public for purposes of § 102(b). *See Egbert v. Lippmann*, 104 U.S. 333, 336 (1881) (holding that a corset worn openly was in public use even though the invention was concealed within the corset); *Hall v. Macneale*, 107 U.S. 90, 97 (1883) (holding that a safe mechanism was in public use even though the invention could not be seen without destroying the safe).

6 *See also Hall*, 107 U.S. at 97; *Egbert*, 104 U.S. at 336; *Netscape*, 295 F.3d at 1319; *New Railhead*, 298 F.3d at 1297; *Lough*, 86 F.3d at 1119; *Sys. Mgmt. Arts*, 87 F. Supp. 2d at 268-70; *3M*, 35 F. Supp. 2d at 1148.

even one use by one person in one government laboratory, is invalidating if, as similar to this case, "co-workers" and "visitors" witnessed the use without a specific obligation of confidentiality. *Id.* at 1058-59.

Similarly, in *Minn. Mining and Mfg. Co. v. Appleton Papers Inc.* ("*3M*"), a district court granted summary judgment of invalidity under § 102(b) due to public use. The *3M* court rejected the patentee's argument that distribution of an invention to the patentee's own employees, to use in connection with their business, was not "public." *3M*, 35 F. Supp. 2d at 1148-49. The court further found that a general duty precluding employees from revealing confidential business information was insufficient, and instead determined that the inventor was required to obtain a specific confidentiality obligation regarding the invention in order to avoid invalidity under § 102(b). *Id.*

In *Netscape Commc'ns Corp. v. Konrad*, the Federal Circuit affirmed summary judgment of invalidity under the "public use" bar of § 102(b), finding that a single demonstration to two people within a university computing lab without any obligation of confidentiality was an invalidating public use. *Netscape*, 295 F.3d at 1323. There, the Federal Circuit rejected the patentee's argument that an implied obligation of confidentiality should be found with respect to uses by the government department that was providing funding for the project. *Id.*

Like the public uses in *Baxter*, *3M*, and *Netscape*, none of MIT's uses of the Back Seat Driver was subject to any obligation of confidentiality. (SOF 8) Like *Baxter*, MIT ([[**REDACTED**]]) Indeed, MIT's policy of ([[**REDACTED**]]) was memorialized in writing (SOF 40, 41), and was followed with respect to MIT's Back Seat Driver.

MIT publicly used and discussed the Back Seat Driver on multiple occasions, each without any confidentiality obligation. (SOF 6-9, 42-57) MIT's Back Seat Driver paper was

presented at the June 6-9, 1989 International Conference on Consumer Electronics, in Rosemont, Illinois. (SOF 43) Neither the conference presentation nor the paper was subject to any confidentiality obligation or restriction. (SOF 43) To the contrary, the paper was made freely available and was provided to all attendees at the conference, as pages 288-289 of a larger collection of presentation materials. (SOF 43-44) This June 1989 Back Seat Driver paper also publicized much about MIT's Back Seat Driver work, including most, if not all, of the aspects of claims 1 and 42 of the '685 patent. (SOF 46) Importantly, the conference paper even mentions the fact that the Back Seat Driver was being used in public all over the Boston area. (SOF 45)

MIT also touted the existence of its Back Seat Driver prototype in a newspaper article, which included inventor interview commentary, a block diagram of the system, and disclosed the subject matter of at least claim 1, including an example of the system's spoken driving instructions. (SOF 48 (July 17, 1989 article in the automotive industry newspaper, *Automotive Electronic News*, which noted "[t]he system, called the Back Seat Driver, gives directions in real time" and that the "prototype guidance system" . . . "uses speech synthesis as a navigation aid.").) These publications alone belie any argument that the public uses, which continued thereafter, were confidential, or that MIT made any effort at all to restrict the flow of information about the project and the public uses.

In addition to publicizing the system prototype and its uses at conference proceedings and in the newspaper, MIT also failed to take any routine confidentiality precautions with respect to the public uses of MIT's Back Seat Driver. (SOF 8, 9, 40, 50-51, 56.) None of the more than fifty people who publicly used the Back Seat Driver before the critical date signed any type of confidentiality agreement. (SOF 9.) When asked to explain what steps MIT took to protect the confidentiality of the Back Seat Driver, Mr. Schmandt testified:

([[**REDACTED**]])

(Tab 40, filed under seal, at 91:2-91:13) However, a mere implied "awareness" of policies not to publish or take credit for the work of others was not sufficient in *Baxter*, and is similarly insufficient for MIT here. *See Baxter,* 88 F.3d at 1059. Likewise, as in *3M*, a general duty not to disclose confidential information is also insufficient. *3M*, 35 F. Supp. 2d at 1148-49.

MIT also gave its sponsor, NEC, complete access to its Back Seat Driver, including demonstrations, with no obligation of confidentiality. (SOF 49-52) Again, Mr. Schmandt ([[**REDACTED**]])

(Tab 40, filed under seal, at 96:22-97:4) But this is insufficient confidentiality for purposes of § 102(b). *See Netscape*, 295 F.3d at 1323 (rejecting patentee's argument that an implied obligation of confidentiality should be found with respect to the public uses by the government sponsor.)

Ultimately, what MIT did with its Back Seat Driver is consistent with MIT's policy of ([[**REDACTED**]]). (SOF 40-41) And Davis' (one of the named inventors) practice regarding other research projects, such as his earlier Direction Assistance project, is also consistent with this policy.[7] There is, thus, no basis for MIT to claim diligence in obtaining or maintaining confidentiality, when the public had extensive, pre-critical-date access to MIT's Back Seat Driver, including fifty public uses, various other demonstrations, conferences, and written materials.

---

[7] Significant aspects of the Back Seat Driver technology were already in the public domain, as early as 1987, by virtue of the publicity surrounding Davis' Direction Assistance system (from which the Back Seat Driver is derived (SOF 62-73)), as well as the public availability of the NEC location system and the DecTalk speech synthesizer. (SOF 71-74) Like the Back Seat Driver, Davis' Direction Assistance project included an interactive system that provided spoken, computer-generated directions for automobile travel within the Boston area. (SOF 62) The system had a telephone interface that used touch tone keypad input and synthetic speech output to provide spoken directions, to successfully direct people throughout Boston. (SOF 64, 67-68) Direction Assistance was in public use in 1987 at the Computer Museum in Boston, and another unit was also in public use as part of the Age of Intelligent Machines exhibit traveling around the United States. (SOF 26)

There is no genuine dispute that each of MIT's fifty public uses was not subject to any obligation of confidentiality. (SOF 8, 9, 43, 48, 50-51, 56-57) Moreover, most, if not all, of the technology involved in the project (particularly the subject matters of claims 1 and 42) was already in the public domain and freely available. (SOF 62-70) As such, there is no genuine issue of fact as to the complete lack of confidentiality surrounding the pre-critical-date uses of MIT's Back Seat Driver.

B. At Least Some of the Pre-Critical-Date Uses of MIT's Back Seat Driver Were For Commercial Purposes.

There is no genuine dispute that at least some of the fifty pre-critical-date uses were for commercial purposes. Evidence of a commercial purpose is strong evidence that a use was a "public use" under § 102(b). Indeed, a use that is commercial in nature can be a "public use" under § 102(b) even if the use is kept in absolute secrecy. *See Kinzenbaw v. Deere & Co.*, 741 F.2d 383, 390 (Fed. Cir. 1984) ("[a] commercial use is a public use even if it is kept secret."). Accordingly, while commercial exploitation is not a necessary requirement for "public use" under § 102(b), a commercial purpose underlying an inventor's use is a strong factor that further supports application of the statutory bar to patentability. *See Invitrogen Corp. v. Biocrest Mfg., L.P.*, 424 F.3d 1374, 1380 (Fed. Cir. 2005) ("[c]ommercial exploitation is a clear indication of public use.").

The undisputed evidence of MIT's commercial purpose for the uses of its Back Seat Driver further bolsters the appropriateness of summary judgment of invalidity under § 102(b). MIT is a corporation and, like most corporations, MIT publishes financial reports each fiscal year. (SOF 58) For Fiscal Year 2006, "Sponsored research" at MIT totaled over a *billion* dollars, representing nearly half of MIT's annual operating budget. (SOF 58) In 1988-89, at the time of the its Back Seat Driver Project, the MIT Media Lab was relatively

new, having only been around since 1985. Nevertheless, the Back Seat Driver project had as its corporate sponsor a subsidiary of the large, Japanese company, NEC, who provided approximately ([[**REDACTED**]]) for the project. (SOF 60) MIT's publication titled "How to Get Value from Media Lab Sponsorships" notes that sponsors of the Media Lab are invited to "visit, view, and discuss" "hundreds of working prototypes developed at the Lab," and which encourages sponsors to "visit the lab during the year for individual discussions and demonstrations." (SOF 59) The same document notes that sponsors are given access to a website that "consolidates technical notes on research projects," and encourages sponsors to "use the Lab as a window to investments and start-ups" in order to get "an inside track on potential opportunities." (SOF 59) This document proves that MIT had a policy of commercially exploiting demonstrations of its research projects as promotional tools to entice Media Lab sponsorship dollars, and that is exactly what MIT did with the Back Seat Driver in 1988-89. One Bellcore witness, indeed MIT's hired expert in this litigation, testified that ([[**REDACTED**]]). (SOF 61) There is no other explanation for the corporate demonstrations. (SOF 7) Accordingly, the commercialization of the Back Seat Driver provides further justification for a grant of summary judgment of invalidity under § 102(b).

II. The Subject Matters Of Claims 1, 42, And 45 Were Embodied In (And Already Reduced To Practice At the Time Of) The Pre-Critical-Date Public Uses

The Court's analysis of the issue addressed in this motion is greatly simplified, because there is no dispute that the subject matters of claims 1, 42 and, at least under MIT's construction, claim 45 were present during the pre-critical-date public uses. Furthermore, the subject matters were not only present, but they had already been "reduced to practice" prior to the pre-critical-date uses. Such reduction to practice precludes, as a matter of law, any experimental use exception to the "public use" bar of § 102(b). *See* Section III. MIT has repeatedly admitted,

both in its interrogatory responses and in contemporaneous documents authored by the named inventors, that claims 1, 42, and 45 were reduced to practice at least as early as June, 1989. *See* Sections IIA, IIB, and IIC *infra; In re Asahi/America, Inc.*, 68 F.3d 442, 444 (Fed. Cir. 1995) (reduction to practice is a question of law). Thus, this court can conclude as a matter of law that claims 1, 42, and 45 were reduced to practice at least as early as June 1989.

A. Claim 1 Was Present and Already Reduced to Practice.

MIT readily admits that the subject matter of claim 1 of the '685 patent was embodied in MIT's Back Seat Driver by April, 1989 (SOF 4, 10), and in every subsequent use. (SOF 12) Thus, it is undisputed that every use after April 1989 included the subject matter recited in claim 1. Furthermore, MIT admits that the subject matter of claim 1 had been reduced to practice at least as early as June 1989. (SOF 15)

MIT also admits that the system (including the subject matter of claim 1) continued to be driven around Boston in July 1989, after claim 1 was already reduced to practice, after Davis already knew the invention would work for its intended purpose (SOF 6, 12), and after Davis had already publicly claimed that it was a "working" system. (SOF 17) Thus, it is undisputed that the subject matter of claim 1 was already reduced to practice prior to pre-critical-date uses of the system.

B. Claim 42 Was Present and Already Reduced to Practice.

MIT twice admitted in its interrogatory responses that the subject matter of claim 42 was reduced to practice "at least as early as June, 1989," noting that "[t]he details of the reduction to practice were fully described in answers to numerous questions to the inventors propounded during the deposition testimony" of the named inventors, and that "those [deposition] answers are herein incorporated by reference." (SOF 28, 29)

Contemporaneous documents undisputedly establish that the subject matter of claim 42 was reduced to practice at least as early as June 1989. (SOF 19-24, 30) As a threshold issue, claim 42 is dependent on claim 1, and thus contains all the limitations of claim 1. However, as explained above, MIT concedes that the subject matter of claim 1 was reduced to practice at least as early as June 1989. (SOF 15) Accordingly, the issue of when claim 42 was reduced to practice is limited to when the additional limitation added in claim 42 (*i.e.*, "wherein each intersection in a route is classified into one type in a taxonomy of intersection types, and the discourse generated in relation to each said intersection depends on its type") was reduced to practice.

A VNIS '89 Back Seat Driver Paper, submitted by MIT in June 1989, admits that the subject matter of claim 42 was embodied in (and already reduced to practice in) Back Seat Driver systems that were being driven around Boston in June and July 1989. (SOF 19-24) The paper itself expressly admits that a prototype had already been "successfully used by drivers who have never driven in Boston," had been running "since April 1989," and included the generation of discourse based on a taxonomy of intersection types, including at least "enter," "exit," and "fork." (SOF 21-24):

- CONTINUE
- FORCED-TURN
- TURN-AROUND
- TURN
- FORK
- ENTER
- EXIT
- ONTO-ROTARY
- EXIT-ROTARY
- STOP

Figure 2: Act taxonomy

(Tab 5, SOF 21)

In fact, this subject matter appears to have been incorporated in MIT's Back Seat Driver from the very beginning, as the Back Seat Driver derived from Davis' earlier Direction Assistance project from 1987. (SOF 25) The concept of classifying intersections in a route into one type in a taxonomy of intersection types, and having the discourse generated in relation to each intersection depend on its type, goes back to Direction Assistance. (SOF 27) A taxonomy of intersection types was in Direction Assistance as early as 1987, including among other intersection types: "enter", "exit", and road "fork." (SOF 27) To provide instructions, Direction Assistance broke the route down into a sequence based upon this taxonomy, and generated sentences accordingly. (SOF 27) In fact, Davis admits that Direction Assistance included a taxonomy of intersection types and generated discourse in relation to each intersection depending on its type. (SOF 27) MIT's own documents admit that the Back Seat Driver was initially based on the Direction Assistance, and that the information contained in the software for Direction Assistance was carried over into the software for the Back Seat Driver. (SOF 25) MIT

further admits that the subject matter of claim 42 was already conceived at least as early as April 1988, at the very beginning of the Back Seat Driver project, further evidencing that it was there from the beginning as a feature carried over from Direction Assistance. (SOF 3) Accordingly, the Direction Assistance evidence alone demonstrates that the additional limitation of dependent claim 42 was present and successfully working as intended in the Back Seat Driver at least as early as June 1989, if not from the very beginning of the project, in April 1988.[8]

Because reduction to practice is ultimately a question of law, which can be resolved here based on MIT's admissions in its (i) interrogatory responses (SOF 28-31), (ii) VNIS '89 Back Seat Driver Paper (SOF 19-24), and (iii) Direction Assistance paper (SOF 27), there is no genuine issue of material fact that the subject matter of claim 42 was present and already reduced to practice at least as early as June 1989 (if not even earlier, such as April 1989), prior to subsequent, pre-critical-date public uses of MIT's Back Seat Driver system.

C. Claim 45 Was Present and Already Reduced to Practice.

Pre-critical-date public uses also embodied the subject matter of claim 45. (SOF 35-38) As with claim 42, MIT twice admitted that the subject matter of claim 45 was reduced to practice "at least as early as June, 1989," citing to both deposition testimony and documents. (SOF 37-39)

---

8 Despite this, on June 16, 2006, the final day of fact discovery, MIT attempted to recant these admissions by amending its response to Interrogatory No. 14 to contend that the subject matter of claims 42 was not reduced to practice until August 4, 1989. (SOF 32) In support of its change of position, the only additional evidence cited by MIT (beyond that already cited in its earlier response, which said claim 42 was reduced to practice "at least as early as June 1989") was Mr. Schmandt's 30(b)(6) deposition testimony. (SOF 32) However, Mr. Schmandt testified that he did not know when the subject matter of claim 42 was first reduced to practice in a working system, because he did not have documents that showed those dates. (SOF 33) Accordingly, Mr. Schmandt's 30(b)(6) testimony does not support MIT's new theory, which MIT raised only after becoming aware of Harman's public use invalidity positions. As described in this Section, the contemporaneous documents (including the very documents cited by MIT in its responses to Interrogatory No. 14) unequivocally prove that the subject matter of claims 42 was reduced to practice at least as early as June 1989, prior to subsequent, pre-critical-date public uses.

These contemporaneous documents establish that, under at least MIT's construction, the subject matter of claim 45 was reduced to practice by June, 1989. (SOF 19-20, 34-39) For example, this subject matter was described in the VNIS '89 Back Seat Driver Paper, camera-ready copies of which were required by June, 1989. (SOF 19, 39, 46) The VNIS '89 Back Seat Driver Paper describes the working embodiment "running in prototype form since April 1989" as giving "instructions just prior to the action" as well as "further in advance, if time permits." (SOF 35) The paper also describes the prototype as giving "the instructions twice, first in detail, and later in a brief form." (SOF 35)

Additional documents cited by MIT in its interrogatory response similarly support its admission of a June, 1989 reduction to practice for claim 45. For example, the June 1989 Back Seat Driver "Abstract" also notes that "[i]f the time between instructions is long, the program gives the instruction twice, first in a detail and later in a brief form." (SOF 38) That same paper notes that "[t]he system has been running in prototype form since April 1989. It has been successfully used by drivers who have never driven in Boston." (SOF 38)[9]

Because reduction to practice is ultimately a question of law, which can be resolved here based on MIT's admissions in its (i) interrogatory responses (SOF 35); and (ii) printed publications, including the VNIS '89 Back Seat Driver Paper and Abstract (SOF 35, 39), there is no genuine issue of material fact that under at least MIT's construction, the subject matter of

---

9 Again, as with claim 42, on June 16, 2006, the final day of fact discovery, MIT attempted to recant its admissions regarding the June, 1989 reduction to practice of claim 45 by amending its response to Interrogatory No. 14 contending that the subject matter of claim 45 was not reduced to practice until August 4, 1989. (SOF 32) In doing so, MIT again cited only to Mr. Schmandt' testimony that he did not know when claim 45 was reduced to practice. (SOF 32-33) As described in this Section, the contemporaneous documents (including the very documents cited by MIT in its responses to Interrogatory No. 14) prove that the subject matter of claim 45 was reduced to practice at least as early as June 1989, prior to subsequent, pre-critical-date public uses.

claim 45 was present and already reduced to practice at least as early as June 1989, prior to subsequent, pre-critical-date public uses of MIT's Back Seat Driver system.

III. As A Matter Of Law, There Can Be No Experimental Use Exception

MIT cannot overcome the § 102(b) statutory bar by arguing that the invalidating use was merely permissible "experimentation" because "experimental use, which means perfecting or completing an invention to the point of determining that it will work for its intended purpose, ends with an actual reduction to practice." *New Railhead Mfg. v. Vermeer Mfg. Co.*, 298 F.3d 1290, 1297-98 (Fed. Cir. 2002); *accord Allen Eng'g Corp. v. Bartell Indus., Inc.*, 299 F.3d 1336, 1354 (Fed. Cir. 2002). Thus, a public use of subject matter that was already reduced to practice is not justifiable as experimental, as a matter of law. *Barmag Barmer Maschinenfabrik AG v. Murata Mach., Ltd.*, 731 F.2d 831, 837-38 (Fed. Cir. 1984); *Pennwalt Corp. v. Akzona Inc.*, 740 F.2d 1573, 1580-81 (Fed. Cir. 1984). As explained in detail in Sections IIA, IIB and IIC above, claims 1, 42, and 45 were already reduced to practice at the time of pre-critical-date public uses, making the experimental use exception unavailable to MIT in this case. To the extent MIT wanted to use its "successful" Back Seat Driver on the public streets of Boston, it was certainly free to do so, but only if it timely filed a patent application that met the requirements of § 102(b). MIT failed to meet this statutory requirement for patentability. Instead, MIT reduced to practice the subject matters of claims 1, 42, and 45, and then continued using and demonstrating the system in public, choosing to delay for more than a year thereafter the filing of its patent application. (SOF 6, 7, 15, 28-31, 36-39) MIT's choice not to follow the statutory mandate of § 102(b) renders claims 1, 42, and 45 invalid as a matter of law.

Nor can MIT argue that its pre-critical-date uses were "experimental" with respect to features of the Back Seat Driver system other than those recited in claims 1, 42, and 45. "Further

refinement of an invention to test additional uses is not the type of experimental use that will negate a public use." *Baxter Int'l., Inc. v. COBE Labs., Inc.*, 88 F.3d 1054, 1060 (Fed. Cir. 1996) (citing *In re Brigance,* 792 F.2d 1103, 1109 (Fed. Cir. 1986) ("experimental use exception does not apply to experiments performed with respect to nonclaimed features") and *Harrington Mfg. Co. v. Powell Mfg. Co.*, 815 F.2d 1478, 1481 (Fed. Cir. 1986) (holding that "testing a tobacco harvester to harvest lower leaves of a tobacco plant was not experimental use where the claim broadly recited removing leaves from part of a tobacco stalk and prior testing had shown efficacy in harvesting upper leaves")). With respect to subject matter recited in claims other than 1, 42, and 45, any alleged experimentation with respect to such subject matter is likewise irrelevant because the statutory bar of § 102(b) is evaluated on a claim-by-claim basis. *See Lough v. Brunswick Corp.*, 86 F.3d 1113, 1122, n.5 (Fed. Cir. 1996) ("[e]ach claim of the patent must be considered individually when evaluating a public use bar."); *Gould Inc. v. United States*, 579 F.2d 571, 581 (Ct. Cl. 1978). Claims 1, 42, and 45 were in public use prior to the critical date and after reduction to practice. No experimental use exception exists under such circumstances. Accordingly, there is no genuine issue of fact that MIT's fifty public uses were not "experimental" and, thus, the claims 1, 42, and 45 are invalid under § 102(b).

IV. Conclusion

For these reasons, Harman respectfully requests that this Court enter summary judgment of invalidity of claims 1, 42, and 45 of the '685 patent under 35 U.S.C. §102(b).

Date: July 25, 2007

Respectfully submitted,

/s/ Courtney A. Clark

Robert J. Muldoon, Jr., BBO# 359480
Courtney A. Clark, BBO# 651381
SHERIN AND LODGEN, LLP
101 Federal Street
Boston, MA 02110

William A. Streff, Jr., P.C.
Michelle A.H. Francis
Craig D. Leavell
Jamal M. Edwards
Colleen M. Garlington
Joanna Belle Gunderson
KIRKLAND & ELLIS LLP
200 East Randolph Drive
Chicago, IL 60601
(312) 861-2000 (phone)
(312) 861-2200 (fax)

Attorneys for Defendant
Harman International Industries, Incorporated

CERTIFICATE OF SERVICE

I hereby certify that this document filed through the ECF system will be sent electronically to the registered participants as identified on the Notice of Electronic Filing and paper copies will be sent to those indicated as non-registered participants on July 25, 2007.

/s/ Courtney A. Clark
Courtney A. Clark

EXHIBIT

1

US005177685A

United States Patent [19]

Davis et al.

[11] Patent Number: 5,177,685

[45] Date of Patent: Jan. 5, 1993

[54] AUTOMOBILE NAVIGATION SYSTEM USING REAL TIME SPOKEN DRIVING INSTRUCTIONS

[75] Inventors: James R. Davis, North Cambridge; Christopher M. Schmandt, Milton, both of Mass.

[73] Assignee: Massachusetts Institute of Technology, Cambridge, Mass.

[21] Appl. No.: 565,274

[22] Filed: Aug. 9, 1990

[51] Int. Cl.[5] G01C 21/00

[52] U.S. Cl. 364/443; 340/988; 364/449; 364/453

[58] Field of Search 340/988, 989, 990, 995; 364/443, 444, 449, 450, 453, 436

[56] References Cited

U.S. PATENT DOCUMENTS

4,139,889	2/1979	Ingels	340/989 X
4,697,281	9/1987	O'Sullivan	455/33 X
4,734,863	3/1988	Honey et al.	340/988 X
4,796,191	1/1989	Honey et al.	364/450
4,812,843	3/1989	Champion, III et al.	340/989 X
4,882,696	11/1989	Nimura et al.	
4,891,761	1/1990	Gray et al.	364/449 X
4,926,336	5/1990	Yamada	364/444

(List continued on next page.)

OTHER PUBLICATIONS

"Synthetic Speech for Real Time Direction-Giving", C. M. Schmandt et al., Digest of Technical Papers, International Conference on Consumer Electronics, Rosemont, Ill., Jun. 6–9, 1989.

"Synthetic Speech for Real Time Direction Giving", C. M. Schmandt et al., IEEE Transactions on Consumer Electronics, 35(3):649–653, Aug. 1989.

"The Back Seat Driver: Real Time Spoken Driving Instructions", J. R. Davis et al., Proceedings of the IEEE Vehicle Navigation and Information Systems Conference, Toronto, Canada, Sep. 1989.

"Back Seat Driver: Voice Assisted Automobile Navigation", by J. R. Davis, Ph.D. Thesis, Massachusetts Institute of Technology, Sep., 1989.

"CD-ROM Assisted Navigation Systems", by O. Ono et al., Digest of Technical Papers, IEEE International Conference on Consumer Electronics, Rosemont, Ill., Jun. 8–10, 1988.

"Attention, Intentions, and the Structure of Discourse", by B. J. Grosz and C. L. Sidner (Computational Linguistics, 12(3):175–204, 1986.

"The Intonational Structuring of Discourse", by J. Hirschberg et al., Proceedings of the Association for Computational Linguistics, 136–144, Jul. 1986.

Automobile Electronic News, vol. 1, No. 16., "U.K. Picks GEC to Head Navigation Project", by James Fallon, Aug. 28, 1989.

"Softening of the Arteries", by Bruce Weber, The New York Times Magazine, p. 78, Aug. 26, 1990.

"Terminal Back Seat Driver", Technology Review, Jul., 1990, p. 10.

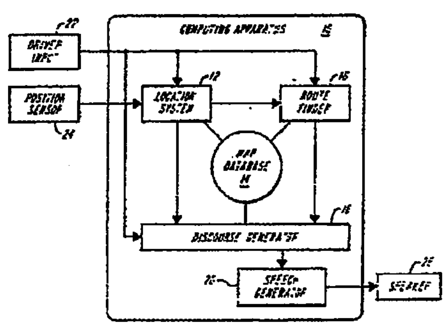
(List continued on next page.)

Primary Examiner—Parshotam S. Lall
Assistant Examiner—Edward Pipala
Attorney, Agent, or Firm—Choate, Hall & Stewart

[57] ABSTRACT

An automobile navigation system which provides spoken instructions to the driver of an automobile to guide the driver along a route is disclosed. The heart of the system is a computing apparatus comprising a map database, route finding algorithms, a vehicle location system, discourse generating programs, and speech generating programs. Driver input means allows the driver to enter information such as a desired destination. The route finding algorithms in the computer apparatus calculate a route to the destination. The vehicle location system accepts input from a position sensor which measures automobile movement (magnitude and direction) continuously, and using this data in conjunction with the map database, determines the position of the automobile. Based on the current position of the automobile and the route, the discourse generating programs compose driving instructions and other messages according to a discourse model in real time as they are needed. The instructions and messages are sent to voice generating apparatus which conveys them to the driver.

58 Claims, 3 Drawing Sheets



U.S. PATENT DOCUMENTS

4,937,751	6/1990	Nimura et al.	364/444 X
4,939,662	7/1990	Nimura et al.	340/990 X
4,951,211	8/1990	De Villeroche	364/444
4,954,958	9/1990	Savage et al.	364/444
4,984,168	1/1991	Neukrichner et al.	364/449
4,992,947	2/1991	Nimura et al.	364/444
5,021,961	6/1991	Ross et al.	340/990 X
5,041,983	8/1991	Nakahara et al.	364/449
5,043,902	8/1991	Yokoyama et al.	364/449

OTHER PUBLICATIONS

"Taxi! Dynamic Cartographic Software for Training Cab Drivers", by M. Bosworth et al., Technical Report, Hunter College Department of Geology and Geography, (212)-772-4000, 1988 paper presented at the Annual Meeting of the Association of American Geographers.

"Digital Map Dependent Functions of Automatic Vehicle Location Systems", by C. B. Harris et al., IEEE Position and Location Symposium, pp. 79-87, 1988, IEEE CH2675-7.

"Digital Maps on Compact Discs", by H. J. G. M. Benning, Technical Paper Series 860125, Society of Automobile Engineers, 1986.

"Eva: An Electronic Traffic Pilot for Motorists", by O. Pilsak, Technical Papers Series 860346, Society of Automotive Engineers, 1986.

"Digital Map Data Bases for Autonomous Vehicle Navigation Systems", by E. P. Neukirchner et al., IEEE Position and Location Symposium, pp. 320-324, 1986, IEEE 86CH2365-5.

"Applications of the Compact Disc in Car Information and Navigation Systems", by M. L. G. Thoone et al., Technical Papers Series 840156, Society of Automotive Engineers, 1984.

"On Board Computer System for Navigation, Orientation, and Route Optimization", by P. Haeussermann, Technical Paper Series 840483, Society of Automotive Engineers, 1984.

"Electro Gryo-Cator: New Inertial Navigation System for Use in Automobiles", by K. Tagami et al., Technical Paper Series 830659, Soc. of Automotive Engineers, 1983.

"Navigation Systems Using gps for Vehicles", by T. Itoh, et al., Technical Paper Series 861360, Society of Automotive Engineers, 1986.

"Extending Low Cost Land Navigation Into Systems Information Distribution and Control", by S. K. Honey et al., IEEE Position and Locations Symposium, pp. 439-444, 1986, IEEE 86CH2365-5.

"Map Matching Augmented Dead Reckoning", by W. B. Zavoli et al., Proceedings of the 35th IEEE Vehicular Technology Conference, pp. 359-444, 1986, IEEE CH2308-5.

"Automated Provision of Navigation Assistance to Drivers", by Matthew McGranaghan et al., The American cartographer 14(2):121-138, 1987.

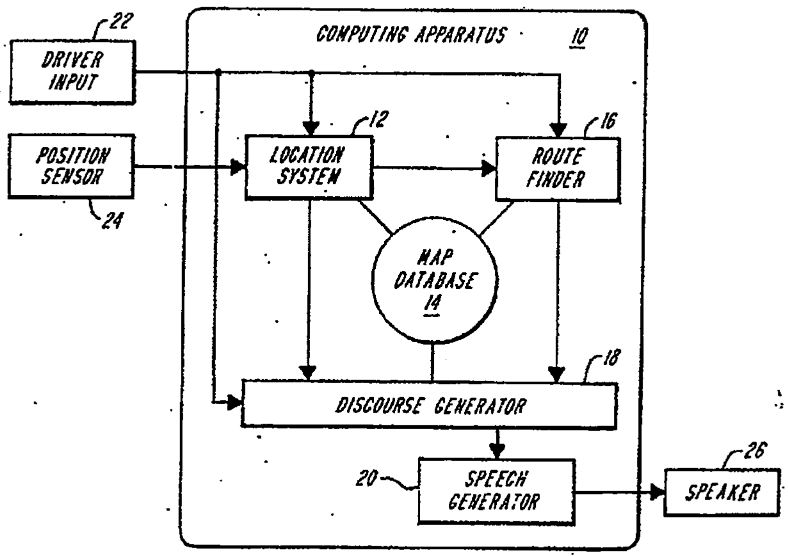

COMPUTING APPARATUS
10
22
DRIVER INPUT
POSITION SENSOR
24
12
LOCATION SYSTEM
16
ROUTE FINDER
MAP DATABASE 14
18
DISCOURSE GENERATOR
20
SPEECH GENERATOR
26
SPEAKER

FIG. 1

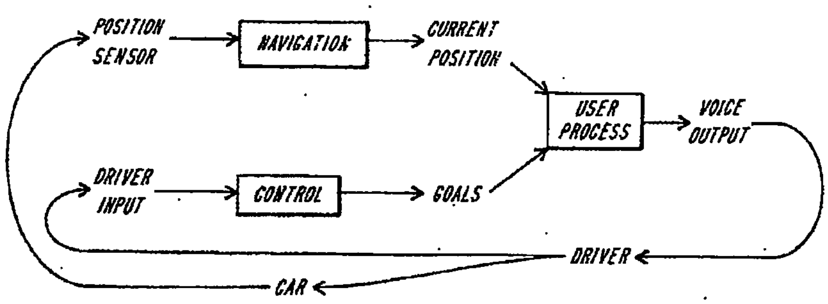

POSITION SENSOR
NAVIGATION
CURRENT POSITION
USER PROCESS
VOICE OUTPUT
DRIVER INPUT
CONTROL
GOALS
DRIVER
CAR

FIG. 2

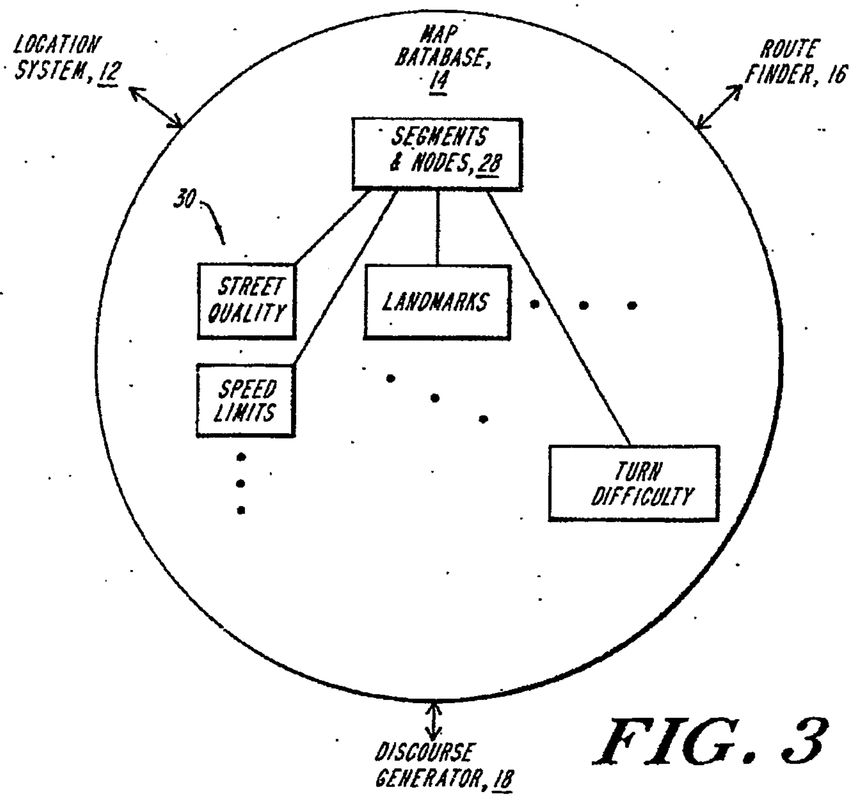

LOCATION SYSTEM, 12
MAP DATABASE, 14
ROUTE FINDER, 16
SEGMENTS & NODES, 28
30
STREET QUALITY
LANDMARKS
SPEED LIMITS
TURN DIFFICULTY
DISCOURSE GENERATOR, 18

FIG. 3

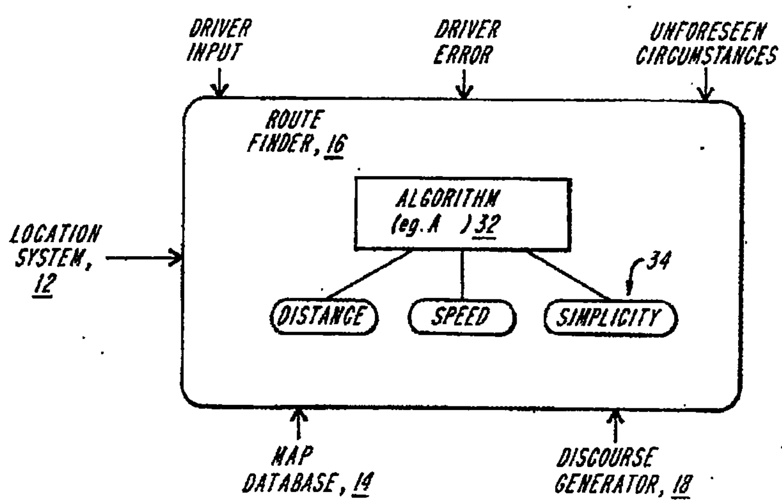

DRIVER INPUT
DRIVER ERROR
UNFORESEEN CIRCUMSTANCES
ROUTE FINDER, 16
ALGORITHM (eg. A) 32
34
LOCATION SYSTEM, 12
DISTANCE
SPEED
SIMPLICITY
MAP DATABASE, 14
DISCOURSE GENERATOR, 18

FIG. 4

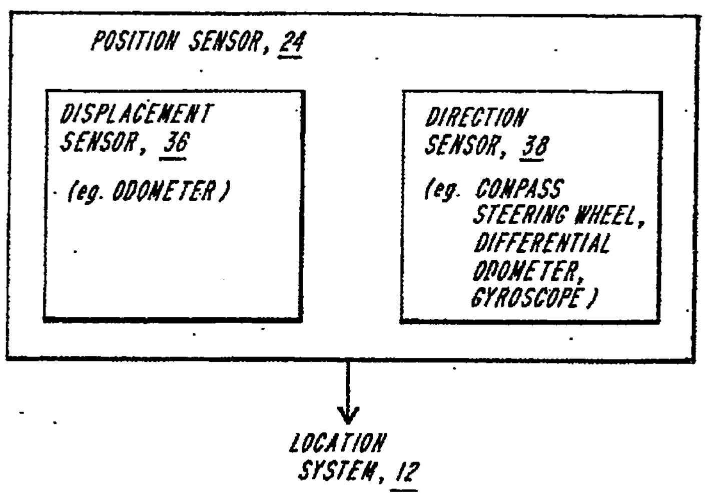

POSITION SENSOR, 24
DISPLACEMENT SENSOR, 36
(eg. ODOMETER)
DIRECTION SENSOR, 38
(eg. COMPASS STEERING WHEEL, DIFFERENTIAL ODOMETER, GYROSCOPE)
LOCATION SYSTEM, 12

FIG. 5

AUTOMOBILE NAVIGATION SYSTEM USING REAL TIME SPOKEN DRIVING INSTRUCTIONS

BACKGROUND OF THE INVENTION

This invention relates to computerized automobile navigation systems, particularly to a system which calculates a route to a destination, tracks automobile location, and provides spoken instructions to the driver in real time as they are needed.

Navigation systems can be classified into three categories:

Positioning systems tell you where you are.

Orienting systems show the direction of your destination.

Instructional systems tell you what to do to get to your destination.

A navigation system can provide one, two, or all of these services. Navigation systems can be further distinguished by how they provide the information:

Verbal systems speak.

Textual systems provide text.

Graphic systems provide pictures.

Finally, systems can be classified as either real time or static. The categories of this classification are not independent. There can be no static positioning system. since one cannot predict the future position of an automobile.

There are several problems with static navigation systems. First, they do not help the driver follow the route. The driver must determine when to apply each instruction. A second problem is that since the instructions must be specified in advance, there is little to be done if the driver does not follow the instructions, which might happen from error, or because the instructions are wrong, or simply ill-advised (as when confronting a traffic jam).

Previous automobile navigation systems have used text or graphics to give navigation information. However, there are several disadvantages to presenting information visually. First, the driver must look at a display while driving, which makes driving less safe. For providing driving directions, visual displays are most easily used when they are least needed. Second, with respect to graphic displays, many people have difficulty using maps, making this mode of providing information undesirable. However, if speech is used, the driver's eyes are left free for driving. In addition, speech uses words, and can therefore refer to past and future actions and objects not yet seen. This is hard to do with symbolic displays or maps.

There is clearly a need for an instructional, verbal, real time automobile navigation system which can guide a driver to a destination much as a passanger familiar with the route would. The present invention meets that need.

SUMMARY OF THE INVENTION

The present invention, called the "Back Seat Driver", is a computer navigation system which gives spoken instructions to the driver of an automobile to guide the driver to a desired destination. Computing apparatus, installed either in the automobile or accessed through a cellular car phone, contains a map database and a route finding algorithm. A vehicle location system uses data from a position sensor installed in the automobile to track the location of the automobile. Discourse generating programs compose driving instructions and other messages which are communicated to the driver using voice generating apparatus as the driver proceeds along the route.

The important differences between The Back Seat Driver and other such systems are that the Back Seat Driver finds routes for the driver, instead of simply displaying position on a map, tells the driver how to follow the route, step by step, instead of just showing the route, and speaks its instructions, instead of displaying them. Each of these design goals has required new features in the programs or in the street map database.

The street map database of the Back Seat Driver distinguishes between physical connectivity (how pieces of pavement connect) and legal connectivity (whether one can legally drive onto a physically connected piece of pavement). Legal connectivity is essential for route finding, and physical connectivity for describing the route.

To find the fastest routes, the map database of the Back Seat Driver includes features that affect speed of travel, including street quality, speed limit, traffic lights and stop signs. To generate directions, the map includes landmarks such as traffic lights and buildings, and additional descriptive information about the street segments, including street type, number of lanes, turn restrictions, street quality, and speed limit. The map also preferably includes other features, such as time-dependent legal connectivity, and expected rate of travel along streets and across intersections. Positions are preferably stored in the map database in three dimensions, not two, with sufficient accuracy that the headings of the streets can be accurately determined from the map segments.

Driving instructions generated by the Back Seat Driver are modeled after those given by people. The two issues for spoken directions are what to say (content) and when to say it (timing). The content of the instructions tells the driver what to do and where to do it. The Back Seat Driver has a large taxonomy of intersection types, and chooses verbs to indicate the kind of intersection and the way of moving through it. The instructions refer to landmarks and timing to tell the driver when to act.

Timing is critical because speech is transient. The Back Seat Driver gives instructions just in time for the driver to take the required action, and thus the driver need not remember the instruction or exert effort looking for the place to act. The Back Seat Driver also gives instructions in advance, if time allows, and the driver may request additional instructions at any time. If the driver makes a mistake, the Back Seat Driver describes the mistake, without casting blame, then finds a new route from the current location.

Giving instructions for following a route requires breaking the route down into a sequence of driving acts, and knowing when an act is obvious to the driver and when it needs to be mentioned. This further requires knowledge about the individual driver, for what is obvious to one may no be so to another. The Back Seat Driver preferably stores knowledge of its users, and uses this knowledge to customize its instructions to the preferences of the users.

Speech, especially synthetic speech, as an output media imposes constraints on the interface. The transient nature of speech requires that utterances be repeatable on demand. The Back Seat Driver has the ability to construct a new utterance with the same intent, but not necessarily the same words, as a previous message.

Synthetic speech being sometimes hard to understand, the Back Seat Driver chooses its words to provide redundancy in its utterances.

An actual working prototype of the Back Seat Driver has been implemented. It has successfully guided drivers unfamiliar with Cambridge, Mass. to their destinations. It is easy to foresee a practical implementation in the future.

BRIEF DESCRIPTION OF THE DRAWING

FIG. 1 illustrates schematically the basic functional components of the Back Seat Driver in its preferred embodiment.

FIG. 2 illustrates the system processes of the preferred embodiment of the Back Seat Driver.

FIG. 3 is a schematic illustration of the map database.

FIG. 4 is a schematic illustration of the route finder.

FIG. 5 is a schematic illustration of the position sensor.

DESCRIPTION OF THE PREFERRED EMBODIMENT

The automobile navigation system according to the present invention is illustrated schematically in FIG. 1. The heart of the system is a computing apparatus 10 comprising a vehicle location system 12, a map database 14, a route finder 16, a discourse generator 18, and a speech generator 20. Driver input means 22 allows the driver to input to the computing apparatus 10 information such as a desired destination. A position sensor 24 measures automobile movement (magnitude and direction) and sends data to the location system 12 which tracks the position of the automobile on the map. The route finder 16 calculates a route to the destination. Based on the current position of the automobile and the route, the discourse generator 18 composes driving instructions and other messages according to a discourse model in real time as they are needed. The instructions and messages are sent to the speech generator 20 which conveys them to the driver by means of a speaker system 26. The speaker system may be that of the car's radio.

In FIG. 1, the computing apparatus is illustrated as a single entity. However, in other embodiments, the components may not all be implemented in the same piece of apparatus. For example, in one working prototype of the Back Seat Driver, the main computing apparatus is a Symbolics Lisp machine, but the location system is implemented separately by an NEC location system that tracks the position of the automobile using its own map database, and the speech generator is implemented separately by a Dectalk speech synthesizer. In another working prototype, the main computing apparatus is a Sun Sparc workstation. The map database for the Back Seat Driver can be provided on a CD-ROM, a floppy disk, or stored in solid-state memory, for example.

The components of the system and the system processes which coordinate their performance, particularly as embodied in the working prototypes, are discussed in the sections which follow. Aspects of the invention are also described in the following sources, which are hereby incorporated by reference:

1. "Synthetic speech for real time direction-giving," by C. M. Schmandt and J. R. Davis, *Digest of Technical Papers, International Conference on Consumer Electronics*, Rosemont, Ill., Jun. 6-9, 1989.
2. "Synthetic speech for real time direction-giving," by C. M. Schmandt and J. R. Davis, *IEEE Transactions on Consumer Electronics*, 35(3): 649-653, August 1989.
3. "The Back Seat Driver: Real time spoken driving instructions," by J. R. Davis and C. M. Schmandt, *Proceedings of the IEEE Vehicle Navigation and Information Systems Conference*, Toronto, Canada, September 1989.
4. "Back Seat Driver: Voice assisted automobile navigation," by J. R. Davis, Ph.D. thesis, Massachusetts Institute of Technology, September 1989.

MAP DATABASE

The map database for the Back Seat Driver in the working prototypes originated as a DIME (Dual Indepedent Map Encoding) file, a map format invented by the U.S. Census Bureau for the 1980 census. Implementing the Back Seat Driver required extending the DIME map format in a number of ways to make it useful for route finding and route describing.

The basic unit of the DIME file is the segment. A segment is a portion of a street (or other linear feature such as a railroad, property line, or shoreline) chosen to be small enough that it is a straight line and has no intersection with any other segment except at its endpoints.

The two endpoints of a segment are designated FROM and TO. If the segment is a street segment (as opposed to, say, a railroad) and has addresses on it, then the FROM endpoint is the one with the lowest address. Otherwise, the endpoint labels are chosen arbitrarily. A segment has two sides, left and right. The sides are chosen with respect to travel from the FROM endpoint to the TO endpoint. A navigator using a DIME file can find the location of an address along the segment by interpolating the addresses between the low and high addresses for the two endpoints. The DIME file is suited to determining the approximate position of a building from its street address.

Attributes of a DIME file segment include: its name (40 characters), its type (a one to four character abbreviation such as "ST"), the ZIP code for each side, and the addresses for each endpoint and each side. At each endpoint of a segment is a pointer to a node. A node represents the coordinates of that endpoint and the set of other segments which are physically connected at that endpoint. Segments share nodes. If any two segments have an endpoint at the same coordinate, they will both use the same node for that endpoint.

A vehicle navigation system using a DIME file can represent the position of a vehicle on the map by a structure called a position. A position has three parts: a segment, an orientation, and a distance. The segment is one of the segments from the map database, the orientation specifies the direction the vehicle is travelling (towards the TO or FROM endpoint), and the distance is the distance from the FROM endpoint of the segment, no matter which way the vehicle is oriented. When travelling towards the TO endpoint of the segment, distance increases, when travelling towards the FROM endpoint, it decreases.

The DIME file is not adequate for routing finding and is only marginal for generating route descriptions. The most important problem with the DIME format is that it indicates only if two segments are physically connected (that is, if they touch), but not whether they are legally connected (i.e. whether it is legal to travel from one to the other). Legal connectivity is crucial for route finding. However, legal connectivity does not

replace physical connectivity; route description requires information about physical connections as well. Physical connectivity also affects route finding directly when seeking the simplest route, since ease of description is determined in part by physical connectivity.

The most significant extension of the DIME file format required for its use in a vehicle navigation system is the explicit representation of legal connectivity. This can be accomplished by adding a legal connection list at each endpoint of a segment to indicate all segments which are legally accessible from that endpoint. This allows the route finder to consider only legal paths. To the inventor's knowledge, this has not been included in any other navigation system.

Another problem with the DIME file is that it is a planar graph. This means that no two segments can cross except at an intersection, so there is no way to correctly represent an overpass, for example. The DIME format represents an overpass by breaking both streets at the point where they cross, and creating a fictitious intersection even though the segments do not touch in reality. These false intersections are particularly troublesome since DIME does not represent legal connectivity, so it appears possbile and legal for a car to jump straight up and turn onto the overpass.

Points in the map database for a vehicle navigation system are therefore preferably three-dimensional. Route descriptions then provide better knowledge of the underlying topography. Stopping distance is affected by slope, so instructions must be given sooner when traveling down a hill. Slope affects safety. The route finder should avoid steep slopes in snowy weather. Finally, distance between points will be more accurate when change in altitude is considered. Roads designed for high speed may be more level than the underlying topography. They may be elevated or they may be depressed. A road which is not at grade will not have the slope of the land beneath it.

Coordinates in the DIME file are stored in ten thousandths of a degree. This means that the position of an endpoint in the map differs form the true position by as much as 6.5 meters in latitude and 5 meters in longitude at the latitude of Boston. This inherent position error causes problems because it introduces error in length and in heading. Uncertainty in heading causes uncertainty in the angle between two segments. A straight street can appear to wobble if it is made of many short segments. Segment "wobble" causes problems for a route finder, makes it hard to generate correct descriptions, and interferes with position determination.

DIME file segment "wobble" can be corrected for by assuming that the angle between two streets is the smallest possible value. However, this sometimes overestimates the speed an intersection can be travelled through. Uncertainty in the angle of segments at an intersection also makes it difficult to describe the intersection correctly and interferes with navigation because it makes it difficult to compare compass headings with the heading of a street.

A richer taxonomy of street types than that provided by DIME is preferable for a vehicle navigation system. Important categories of streets are: ordinary street, rotary, access ramp, underpass, tunnel, and bridge. Preferably, non-streets such as railroad, water, alley and walkway are also included.

The DIME file records a small amount of information about each segment. For a vehicle navigation system, additional attributes are preferably added to make better descriptions. Important additional attributes are street quality, divided roads, signs, traffic lights, stop signs, buildings, other landmarks, lane information, and speed limit.

The street quality can be, for example, a number from 1 ("super") to 4 ("bad") which combines the ease of locating and following the street and the expected rate of travel along it. The street quality attribute should be used by both the route finder and the route describer.

The identification of divided roads is necessary to avoid U Turns where they are not possible, although it is preferable to make U Turns only if there is no other alternative. In addition, the route finder should recognize that a divided road is safer than an undivided road.

Sign and exit numbers are preferably stored in the map database as connection cues, which are text strings that give cues for moving from one segment to another. Every cue has a type which tells the kind of cue, e.g. sign or exit-number. There may be more than one connection cue for a given pair of segments, but there should never be more than one of a type.

The most useful landmarks are traffic lights. Traffic lights are preferably stored independently for each endpoint of each segment, since the presence of a light at one segment of an intersection does not imply that all other segments at the intersection have a light.

Two types of buildings which are especially useful as landmarks are toll booths and gas stations. Toll booths can be stored as connection cues. Gas stations can be stored in the services database described below. However, a preferred approach is to index gas stations (and other buildings) by street.

Roads often have more than one lane. Selecting the proper lane can make travel faster, and it may even be mandatory, since certain turns may only be possible from some lanes. The map database therefore preferably contains the number of lanes for both directions on a segment, and whether one or more lanes is reserved for turn restrictions.

The map database also preferably includes time dependent legal connectivity. Sometimes a given turn will be prohibited at certain hours of the day, typically rush hour. Additionally, lanes sometimes switch direction during the day to accomodate rush hour traffic, and some lanes are reserved for carpools during rush hour.

The expected rate of travel is not necessarily a function of street quality. Although there is a correlation, travel rate is preferably a separate segment attribute. One reason is that travel rate, unlike quality, changes during the day. A model of traffic flow like that of an experienced driver (i.e. it should know what "rush hour" means) is preferably implemented in the map database.

Some turns, though legal, are difficult to make. The route finder preferably avoids these turns if possible. To an extent, the difficulty of a turn is implicit in the quality of the participating street segments, but an explicit model in the map database is preferred.

Some lanes or streets are restricted to certain kinds of traffic (car pools, no commerical vehicles). Also important are height restrictions, since some underpasses are so low that tall vehicles will not fit under them. This information is preferably included in the map database.

At some lights it is permitted to make a right turn at a red light after a full stop. Right turns here will be no slower than rights turns at a stop sign, so the route finder should prefer such intersections to those that do not permit it. Also, traffic lights have differing cycle

lengths. The map database preferably includes this information.

Local knowledge is also preferably included in the map database. These are facts about how people and institutions act on or near the road; e.g. that a speed trap is here, or that this road is one of the first ones plowed after a snow storm.

The Back Seat Driver should allow the driver to select famous destinations by name in addition to address by including this information in a database, and this database should be integrated with the services database, discussed below. The Back Seat Driver should also support names of buildings and office plazas made up by developers without reference to the street names.

Service locations are preferably stored in a services database. This database lists services such as gas stations, automatic teller machines and stores. For each service is recorded the name of the establishment, the address, phone number, and hours of operation. This allows the Back Seat Driver to select the closest provider of a service known to be open. The database can also be used as a source of landmarks when giving directions.

The map database preferably contains information on the division of the city into neighborhoods. This is useful for selecting an address. The postal ZIP code is not good for classifying neighborhoods.

Pronunciation information is preferably stored in a database for those place names which are easily mispronounced by the speech synthesizer. It would also be desirable to record which of those names have unusual spellings. This would allow the system to warn the driver to be alert for signs that might otherwise surprise her. Note that the driver only hears the name of a street, and has to guess how it is spelled from the sound she hears.

Abbreviations are preferably included to allow the user to enter certain street names in abbreviated form. A second use for abbreviations is to supply alternate spellings for streets, for example, to allow the driver to spell "Mt Auburn" as "Mount Auburn".

An almanac is preferably included to list the time of sunrise and sunset for the city. Arrangements can be made to either purchase this database or locate a program which an calculate it, for arbitrary position and date.

A problem for a practical Back Seat Driver is how to keep the map database accurate, since the streets network is constantly changing. Over time, new street are constructed, old streets are renamed or closed. These kinds of changes are predictable, slow, and long lasting. Other changes are unpredictable, quick, and transient. A road may be closed for repairs for the day, blocked by a fallen tree, or full of snow. Such changes are usually short lived. Thus, the Back Seat Driver needs the ability to change legal connectivity dynamically. In addition, the route finder should preferably have the ability to avoid congested roads caused by rush hour or accidents, for example. The map database is therefore preferably continuously updated by some form of radio broadcast by an agency that monitors construction and real time traffic conditions.

The Census Bureau, in cooperation with the United States Geological Survey, has designed a new map format known as TIGER (Topologically Integrated Geographic Encoding and Referencing) which has several improvements over the DIME format, but which is still a planar graph representing only physical connectivity. The map database for a Back Seat Driver could be also be originated from a TIGER file as long as the extensions discussed above were implemented.

The map database is shown schematically in FIG. 3. In the preferred embodiment, the map database 14 includes, as its basis, a file 28 of segments and nodes. File 28 may be an original file or may be adapted from a DIME file or a TIGER file by adding the above-described extensions. In addition, the map database 14 may include optional features 30, as described above.

ROUTE FINDER

Finding a route between two points in a street network is an example of searching a general graph. The task is to find a sequence of segments that lead from the origin to the destination. There are usually a great many distinct ways of getting from one place in the city to another, some better than others. Graph search algorithms differ in the quality of the solution they find and the time they require. The Back Seat Driver requires an algorithm that finds a good route in a short time.

The route finder of the working prototypes of the Back Seat Driver is based on an A* search algorithm. The A* algorithm is a form of best-first search, which itself is a form of breadth-first search. These searching techniques are well-known and are described in detail in Davis, 1989, cited above.

In a breadth-first search, a tree of all possible decisions is divided into levels, where the first level actions are those leading from the root, the second level actions are those that come from situations after first level actions, and so on. All actions at a given level are considered before any at the next higher level. While the breadth-first search is operating, it maintains a list of all possible partial routes and systematically examines every possible path from the end of every partial route to compile a new list of partial routes. This search procedure finds the path with the fewest segments. However, this is not necessarily the best path. To be sure of finding the best path, the search cannot stop when the first path is found, but must continue, expanding each path, until all paths are complete. This is not at all desirable, since there could be (and in fact will be) many paths.

The best-first algorithm solves this problem by keeping track of the (partial) cost of each path, and examining the one with the smallest cost so far. This requires a function that can compare two routes and produce a numeric rating. Such a function is called a metric. To further reduce the cost of searching, before adding a segment to a path, the best-first search checks to see whether it is a member of any other path. If it is, it is not added, for presence on the other path means that the other path was a less expensive way of reaching the same segment.

Best-first search finds the best solution and requires less time than exhaustive breadth-first search, but it still must consider partial solutions with an initial low cost which prove expensive when complete. The A* algorithm avoids wasting time on such falsely promising solutions by including an estimate for the completed cost when selecting the next partial solution to work on. The cost estimate function is $f^*(r)=g^*(r)+h^*(r)$, where r is a route, $g^*(r)$ is the known cost of the partial route, and $h^*(r)$ is the estimate of the cost to go from the endpoint of the route to the goal. The h^* function must have the property of being always non-negative and

never over-estimating the remaining cost. An h* meeting these two conditions is said to be admissible. It should be obvious that if h* is chosen to be always zero, then A* search is just best-first search. In applying A* to finding routes on a map, h* is just the cartesian distance between the endpoint of the partial route and the destination point. It is certain that no route will be shorter than the straight line, so this estimate is never an over estimate. A* search is more efficient than best-first.

The A* algorithm finds the optimum route, but the Back Seat Driver might be better served with an algorithm that finds a reasonable route in less time. This is especially true when the vehicle is in motion. The longer the route finder takes, the greater the distance that must be reserved for route finding. As this distance becomes larger, it becomes harder to predict the future position of the car. This can be done by choosing an h* which multiplies the estimated distance remaining by a constant D. Setting D greater than one makes h* no longer admissible, since the estimate might exceed the actual cost by a factor of D. The resulting routes are no longer optimal. but are still pretty good. The effect is to make the algorithm reluctant to consider routes which initially lead away from the goal.

The route finder preferably uses a value of 2 for D. This yields the greatest increase in payoff. A possible improvement is to run the route finder twice, first with a high value of D to find an initial route in order to begin the trip, and then with a low D to search for a better route, using spare time while driving.

Preferably. three different metrics are used. The distance metric finds the shortest route, the speed metric finds the fastest route, and the ease metric finds the easiest route. The metric for distance is just the sum of the lengths of the component segments. The other two metrics are more complicated than the distance metric. because they must consider intersections as well as segments. In general there is a cost to travel along a segment and a cost to get from one segment to another. All costs are expressed as an "equivalent distance" which is the extra distance one would travel to avoid the cost.

The metric for speed estimates the cost for traveling along a segment by multiplying its length by a constant which depends upon the quality of the street. In principle, one could calculate expected time by dividing length by the average speed on the segment were this quantity available in the database. Examples of appropriate constants are:

Quality	Factor
super	1
good	1.2
average	1.5
bad	2.0

All multiplicative constants must be greater than or equal to one to ensure that the cost of a route is never less than the straight line distance between two points. This condition is essential for the correct operation of the A* search algorithm, since the estimation function (g*) must always return an under-estimate.

The time to cross an intersection is preferably modeled by a mileage penalty which depends upon the nature of the intersection. Examples of appropriate penalties are:

Factor	Cost	Reason
turn	[illegible] mile	Must slow down to turn
left turn	[illegible] mile	May have to wait for turn across traffic flow
traffic light	[illegible] mile	Might be red

If the segment is one-way, the penalties should be cut in half, since there will be no opposing traffic flow. The turning penalties should be computed based only on the angle between two segments, not on the segment type or quality.

The metric for ease seeks to minimize the driver's effort in following the route. Again, driver's effort is the sum of the effort to travel along a segment and the effort to get from one segment to another. Travel along a segment depends upon its quality. Turns of every sort should be penalized equally, since they all require decisions. The intention of this metric is to find routes which require the least amount of speaking by the Back Seat Driver, leaving the driver free to concentrate on other matters.

If the driver leaves the route, the Back Seat Driver must immediately inform the driver and begin to plan a new route. Route planning after a mistake is no different from any other time, except that the vehicle is more likely to be moving. In the working prototypes, when the car is moving. the Back Seat Driver first estimates the distance the car will travel during the route finding process by multiplying the current velocity by the estimated time to find the route. Then it finds the position the driver will reach after traveling this distance, assuming that the driver will not make any turns without being told to do so. It then finds a route from this extrapolated position to the goal. Finally, it finds a route from the car's actual position to the estimated starting position. This second route is so short that the car is unlikely to move far during the time it is computed.

The route finder of the working prototypes estimates the time to find the route between two points by multiplying the distance between them by a constant. This constant was initially determined by running the route finder for 20 randomly selected pairs of origins and destinations. As the Back Seat Driver runs, it accumulates additional values for the constant.

A problem is how to reliably detect when the driver has left the route. With the extended DIME format of the working prototypes, if the driver turns into a gas station, for example, the system will believe, falsely, that the driver has turned onto some street, because the street map includes only streets, and not other paved areas such as parking lots and filling stations. From this false belief, the system will conclude that the driver has made a mistake. However, this problem can be solved by increasing the detail of the map.

Sometimes the driver will choose to not follow a route for good reasons that the Back Seat Driver is unaware of, perhaps because the road is blocked or because of a traffic jam. For the first case, the driver should be provided an "I Can't Do It" button or other means to inform the Back Seat Driver that the road is (temporarily) blocked. Once informed, the Back Seat Driver must automatically find a new route. For the second case, the driver's only recourse is to cancel the current trip (by pushing another button, for example), and, once out of the situation, re-request a route to the original destination. It is essential. though, that the

driver either notify the Back Seat Driver of the impossibility of the requested action or cancel the trip, because otherwise the Back Seat Driver will treat the deviation from the route as a mistake, and continue to attempt to find a new route, which may very well lead back through the street the driver is trying to avoid.

The route finder is shown schematically in FIG. 4. In the preferred embodiment, the route finder 16 includes, as its basis, an algorithm 32. Algorithm 32 may be, for example, an original algorithm based on a best-first search algorithm the A* algorithm, or a modified A* algorithm. In preferred embodiments, the route finder is adapted to find the best route according to any one of three cost metrics 34: distance, speed, simplicity. The route finder calculates a new route in the event of driver error or unforeseen circumstances, as indicated.

LOCATION SYSTEM AND POSITION SENSOR

The Back Seat Driver must know the position of the vehicle. This can be achieved using available technology adapted for the requirements of the Back Seat Driver. At a minimum, the location system for a vehicle navigation system must determine the vehicle position to the nearest block. If it is to tell the driver when to turn, it must be able to distinguish between the closest of two adjacent turns.

Consideration of the Boston street map shows that it has many streets which are both short and a possible choice point. Based on a study of the percentage of segments which are shorter than a given length, an accuracy of 10 meters is desirable. This is a higher accuracy than has been specified in prior art approaches (see Davis, 1989, cited above). The Back Seat Driver's use of speech imposes strict requirements on position because of limitations on time. Unlike a display, speech is transient. An action described too soon may be forgotten. The Back Seat Driver is intended to speak at the latest time that still permits the driver to act. Allowing two seconds for speech, a car at 30 mph covers 27 meters. This consideration suggests a minimum accuracy of 15 meters.

Location systems can be divided into two categories:

Position finding systems determine position directly by detecting an external signal.

Position keeping (dead reckoning) systems estimate the current position from knowledge of an earlier position and the change in position since that position.

All existing position finding systems use radio signals. The broadcast stations may be located on street corners, seacoasts, or in orbit around the earth. Systems differ in their range, accuracy, and cost. A survey of those systems which might plausibly be used for automobile navigation is included in Davis, 1989, cited above.

Position keeping (dead reckoning) systems obtain position indirectly, by keeping track of the displacement from an originally known position. One can measure displacement directly, or measure velocity or acceleration and integrate over time to obtain displacement.

The forward motion of a car is measured by the odometer. On late model cars, the odometer cable has been standardized. It revolves once every 1.56 meters. This is a reliable measure of forward progress, as long as the wheels do not slip. Measuring direction, though, is more difficult. Among the possibilities are:

magnetic compass A magnetic compass has the advantages of small size and ease of use, but is unreliable because of variation between magnetic and true north and deviation due to the ferrous material of the car and magnetic flux arising from electric currents within the car.

steering wheel The steering wheel could be instrumented to measure the amount of turning.

differential odometer When a car turns, the two rear wheels travel different distances, and thus rotate at different rates. Measuring the difference in rotation provides an indication of amount of turning. This differential rate of rotation is just what is measured by anti-skid brakes, so no additional instrumentation is required to obtain this measure for an automobile suitably equiped.

gyroscope Gyroscopes measure angular acceleration. The familiar rotation gyroscope and esoteric laser ring gyroscope are not suitable for automotive use because they are too expensive. Lower cost alternatives are the rate gyro and the gas jet gyro. The rate gyro measures angular acceleration in a vibrating piezo-electric substance. The gas gyro measures turn (or yaw) rate. In this design, a jet of gas travels down the center of a sealed tube. Anemometers are placed on either side of stream. When the car turns, the stream is deflected and the velocity is measured. The velocity of the gas at the anemometer is proportional to the turn rate. Such devices can measure turn rates of up to 100 degrees per second, with a noise of about one half degree/second.

The position sensor is shown schematically in FIG. 5. As indicated, it includes a displacement sensor 36 and a direction sensor 38.

A position keeping system with no error could be calibrated when installed, and then maintain its own position indefinitely. Unfortunately, errors arise in measuring both distance and heading. Sources for error include difference in tire pressure, composition and wear, slipping, cross steering from winds, change in tire contact path in turns, magnetic anomalies, and gyro noise. The NEC dead reckoning system, employed in the prototypes of the Back Seat Driver, accumulates about one meter of error for every ten meters traveled. The error is even worse when traveling near railroads because the NEC system uses a magnetic compass.

Some dead reckoning systems recalibrate themselves to eliminate systematic errors. Such recalibration is possible when the vehicle is at a known position or when stopped. One way to correct dead reckoning errors is to use knowledge of the road network as a constraint on position, in what is known as map matching. Map matching requires that the position keeping system have a map of the locale where the vehicle is being driven, and is based on the assumption that the vehicle is always on a street present in the map. If the estimated position falls to one side of the road, the estimate can be corrected. When the vehicle makes a turn, the system assumes the vehicle is at the closest intersection, and thus the absolute position can be estimated. Every dead reckoning system uses some form of map matching. Map matching reduces the stringency of position keeping, but accuracy remains a concern, since the system must maintain its position when the driver drives off the map, e.g. into a driveway or a parking lot.

In the working prototypes, a unit built by NEC Home Electronics, Ltd. is employed. It is a dead-reckoning position keeping system which uses speed and direction sensors. To compensate for error, it uses map matching on a map database stored on CD-ROM. The unit is described in "CD-ROM Assisted Navigation

Systems" by O. Ono; H. Ooe, and M. Sakamoto, in *Digest of Technical Papers, IEEE International Conference on Consumer Electronics*, Rosemont, Ill., Jun. 8–10, 1988.

As implemented in the working prototypes, the map database used by the location system is completely distinct from the map database used by the route finder and discourse generator. This is unfortunate since the maps will not always agree unless they are kept equally up-to-date. However, in other embodiments within the scope of the invention, the location system uses the computing resources and map database of the main computing apparatus illustrated in FIG. 1. Positioning systems for the Back Seat Driver preferably combine position keeping and position finding, since neither alone will work all the time. A position keeping system needs periodic corrections, but a position finding system that depends on radio reception will not work in tunnels or bridges. Hybrid systems which could be used by the Back Seat Driver are referenced and discussed in Davis, 1989, cited above.

DISCOURSE GENERATOR

The Back Seat Driver attempts to provide instructions to the driver as a passenger in the car familiar with the route would. The content and timing of the instructions and other messages described below are based on a study of natural driving instruction described in detail in Davis, 1989, cited above.

To the Back Seat Driver, a route is a sequence of street segments leading from the origin to the destination. Each connection from one segment to another is considered an intersection, even if there is only one next segment at the intersection. At any moment, the car will be on one of the segments of the route, approaching an intersection. The task of the Back Seat Driver is to say whatever is necessary to get the driver to go from the current segment, across the intersection, to the next segment of the route. Most often, nothing need be said. But at other times, the Back Seat Driver will need to give an instruction.

Instructions must use terms familiar to the driver. An example is what to say at a fork in the road. Considering only topology, there is no difference between a fork and a turn, but it would be confusing to call a fork a turn.

The two key issues in describing a route are deciding what to say and deciding when to say it. There is a tradeoff between these two factors. At one extreme are directions given completely in advance, with no control over when the driver reads them. A direction of this kind might be: "Go half a mile, then take a left onto Mulberry Street". A driver following such an instruction must use the odometer to estimate distance or look for a street sign. The instruction itself does not say when to act. The other extreme are instructions which rely totally on timing for success. Such an instruction might be: "Turn left now".

An intersection type is called an act because the important thing about an intersection is what action the driver takes to get across it. The Back Seat Driver is preferably implemented with an object-oriented programming methodology, so for each act there is an expert (an object) capable of recognizing and describing the act. The Back Seat Driver generates speech by consulting these experts. At any moment, there will be exactly one expert in charge of telling the driver what to do. To select this expert, the Back Seat Driver asks each expert in turn to decide whether it applies to the intersection. The experts are consulted in a fixed order, the most specific ones first. The first expert to claim responsibility is selected. This expert then has the responsibility of deciding what (if anything) to say.

Each act has a recognition predicate which is used to determine if a given intersection should be classified as that act. A predicate can consider topology, geometry, the types of street involved, or any other factor. The predicate also decides whether the move is obvious, that is, the driver can be trusted to do it without being explicitly told to do so. Actions that are obvious are not described. If the next action is obvious, the Back Seat Driver looks ahead along the route until it finds one which is not obvious. There will always be at least one, because stopping at the end is never obvious.

The acts in the working prototypes include CONTINUE, FORCED-TURN, U-TURN, ENTER, EXIT, ONTO-ROTARY, EXIT-ROTARY, STAY-ON-ROTARY, FORK, TURN and STOP.

A CONTINUE is recognized when the driver is to stay on the "same" road. Almost always, a continue is obvious and nothing should be said. The continuation of a street depends on the type of street: from a rotary, it is the next rotary segment; from an access ramp, if there is exactly one next segment, that is the continuation, otherwise there is no obvious next segment; otherwise, it is the one segment that requires no more than 30 degrees of angle change (if there is exactly one, and if it is not a rotary) or the one segment with the same name (if there is exactly one). The reason for comparing names is not because the driver is aware of the name, but because the designer who named the street was. The assumption is that if two segments have the same name, they are the same street, and that is why they have the same name. This "sameness" is presumably reflected in details not captured by the map, for example continuity of painted centerline. There are many places in the area where the obvious "straight" continuation of a segment is at an angle as great as 45 degrees, but it would not be right to call this a turn.

A FORCED-TURN is an intersection where there is only one next street segment where the road bends more than 10 degrees. Even though there is no decision to make at a forced turn, it is useful to mention because it strengthens the driver's sense that the Back Seat Driver really knows about the road conditions. A forced turn is not worth mentioning if both segments are part of a bridge, a tunnel, or an access ramp, or if the angle is less than 20 degrees.

The U-TURN action is recognized when the heading of the car is the opposite of what it should be. Recall that a route is a sequence of segments and endpoints. At all times the car will be on one of the segments in the sequence. If the car's orientation is not the same as the endpoint in the path, then the driver must turn around. Preferably, the route finder only calls for a U Turn if there is no other way.

To ENTER is to move onto a super street (or an access ramp that leads eventually to a super street) from an ordinary street, but not from a super street or an earlier access ramp. Similarly, to EXIT is to move from a super street onto a street with lesser quality that is either an access ramp or has a different name. Some super streets are not uniformly super and it would not be right to call the change in quality an exit.

To go ONTO-ROTARY, to STAY-ON-ROTARY, and to EXIT-ROTARY are acts which can be correctly

described only if the street map database includes an explicit marking of streets as rotaries.

At a FORK, there must be at least two alternatives, all within a narrow angle, and none of the branches must be the obvious next segment—that is, the branches must all be more or less equal. Either all the alternatives must be access ramps, or none of them must be. A branch can only be considered obvious if it is the only branch with the same level of quality, or if it is markedly straighter than the others, or if it is the only one with the same number of lanes, provided that all of these clues agree. If one branch is stronger than the others, the intersection is not a fork. It is either a continue or a turn.

The STOP action is recognized when the vehicle is on the destination segment. Finally, a TURN is an intersection not handled by one of the above cases. The greatest weakness of the above approach is that the recognition predicates are sensitive to small changes in the angles between segments. It is not likely that people use absolute numbers (e.g. 10 degrees) as cut-off values for their determination of how to describe an intersection. More likely, different classifications compete. Still more important, people making classifications use visual cues, not just facts from the map.

Each act has a description function to generate a description of the action. The description function takes inputs specifying the size of the description (brief or long), the tense (past, present or future), and the reference position. A short description is the minimum necessary for the act. It is typically an imperative (e.g. "Bear left."). A long description includes other facts about the action, an expression indicating the distance or time until the act is to be performed, and possibly information about the next act, if it is close. The reference position is a position (along the route) from which the action is to be described.

A brief description consists only of a verb phrase. The verb depends on the type of act and perhaps on the specifics of the act. Besides the verb itself, the verb phrase must say which way to go. In most cases, the word "left" or "right" is sufficient. Where it is not, the possibilities are to use a landmark or to describe the turn. A landmark can be either in the appropriate direction ("towards the underpass") or the other direction ("away from the river"). Specifying direction with a landmark has the advantage that some drivers confuse left and right sides, or mishear the words, so it is a redundant cue. Also, it increases the driver's confidence that the system really knows what the land looks like. A description of the turn can mention either quality or the relative angle of the desired road. The angle must be described qualitatively (more or less "sharp"). It would be more precise to use the angular distance (e.g. "turn right 83 degrees"), but drivers would not understand it. Preferably, the expert for each act follows a protocol which includes:

recognize?—is a proposed movement an example of this kind of driving act?

instruction-vp—generate a verb phrase describing this act

instruction-np—generate a noun phrase describing the act

position-to-doit—the position where the driver would begin carrying out the act

obvious?—would the driver do this act without being told?

sentences—generate all sentences needed to describe this act

congratulate?—should the driver be congratulated after carrying out this kind of act

The following sample is a Back Seat Driver description of the left turn from Fulkerson Street to Main Street in Kendall Square, Cambridge, Mass.:

> Get in the left lane because you're going to take a left at the next set of lights. It's a complicated intersection because there are two streets on the left. You want the sharper of the two. It's also the better of them. After the turn, get into the right lane.

This instruction begins with a piece of lane advice, an action to be taken immediately, then describes an action in the near future. The action is a turn, though that word is not used explicitly. It tells the direction of the turn (left) and specifies a landmark (the lights) that says where the turn is. In many cases, this would be enough, but here there are two streets on the left, so the instruction goes on to specify the desired road in two ways (by comparative position and relative quality). Finally, it concludes with some lane advice to be executed during (or just after) the act.

The above example is the most complicated text that the Back Seat Driver prototypes have produced. Length and detail are not virtues in giving directions. The Back Seat Driver produces a text this long only because it does not have better means to make the driver follow the route. If a shorter text would accomplish the same aim, it would be better.

Besides telling drivers what to do, the Back Seat Driver must also tell them when to do it. One way to do this is by speaking at the moment to act, but it is useful to also give instructions before the act, if time permits. This allows time for preparation, if required, permits the driver to hear the instruction twice, and also spares the driver the need to be constantly alert for a command which must be obeyed at once.

When an act is more than a few seconds in the future, The Back Seat Driver uses a long description, which includes one or more cues which either describe the place for the act, the features of the road between the current location and the place, or the distance or time until the act. This description should be so clear that the driver cannot only recognize the place when it comes, but can also be confident in advance that she will be able to recognize the place.

The Back Seat Driver preferably uses a landmark as a cue when it can. A numeric distance is the cue of last resort. However, some drivers prefer to also hear distances, especially if the distance exceeds a certain threshold. Therefore, a parameter is preferably included in the user-model, described below, for this minimum distance expressed as a number. If the distance is below this, a qualitative phrase is produced by the discourse generator, if above, a number is produced. The cutoff can be zero, in which case numbers are always used, or set to an infinite value, in which case they never are.

A cue is expressed either as a full sentence ("Drive to the end of the street, then . . . ") or a preposed preposition phrase ("At the next set of lights, . . . "). Research has shown that a cue should not be expressed by a preposition after the verb as in "Take a left at the lights." because some drivers start to take the left as soon as they hear the word "left". This may be because syn-

thetic speech does not provide enough intonational cues for the driver to reliably predict the length of the sentence, leading the driver to act on syntactic information alone, and thus taking the sentence to be complete as soon as the word "left" is heard.

The description of a road feature depends upon whether or not it is visible. If it is, it can be referred to with a definite article ("the rotary", "the overpass"). If not, an indefinite article is used. The program cannot tell whether an entity is actually visible, so it uses distance as an approximation. If the feature is closer than one tenth of a mile, it is considered to be visible.

A special case of cues is when the driver is at the place to act. When stopped a few meters from the intersection, it is wrong to say "Turn at the next lights" even if it is literally true. In the working prototypes, the Back Seat Driver thinks of itself as being at that intersection it if its less than thirty yards away, except that if there is a stop light at the intersection and the car is not moving, then the intersection distance if fifty yards, since cars might be backed up at such an intersection. When at an intersection, the Back Seat Driver should say "Take a left here" rather than "Take a left now" because drivers waiting for a traffic light will rightly resent being told to do something they have good reason not to do.

Traffic lights are very good landmarks because they are designed to be easily seen and drivers have an independent reason to watch for them, namely a desire to avoid accidents. When referring to a traffic light, if the car is at the intersection for the lights, the Back Seat Driver should use a proximal deictic ("this" or "these", as opposed to the distal "that" or "those") to show it means the lights that are here.

The Back Seat Driver preferably has a database of buildings as part of its directory of services. If it finds a building on the corner, it should include it as a potential landmark: e.g. "Look for Merit Gas on the left side".

Other landmarks are features of the road, such as underpasses: bridges, tunnels, bends in the road, and railroad crossings. Still another potential landmark is the road coming to an end. This is a landmark that is impossible to miss. However, research has shown that if the Back Seat Driver says "Drive all the way to the end, then . . . ;" some drivers take "the end" to mean not "the farthest you can go along this road" but just "the next intersection".

A street name can be a landmark, but not a good one, unless the driver already knows the street. There are several reasons why street names should not be used. First, the driver may not hear the name correctly. Second, the driver may hear the name, but not know how to spell the name after hearing it, so she may not recognize the name in its printed form. This is especially a problem when the driver is from out of town. Finally, even if the driver knows the spelling, street signs are often missing, turned around, or invisible due to weather or darkness. Despite all the problems that come with using street names, many drivers ask for them. To accommodate them, a parameter in the user-model is preferably included to control the use of names. If set, names are supplied as part of the instruction. When names are included, they are preferably attached at the end of the instruction ("Take the second left. It's Elm Street.") rather than directly ("Take the second left onto Elm Street."), which weakens their salience somewhat, and makes them more of a confirmatory cue than an essential one.

Signs can be important landmarks. A problem with using signs as cues occurs, however, if the information in the sign is stored as unstructured text in the map database. It is important that the Back Seat Driver understand what the sign says, not simply utter the words. There are two reasons for this. First, the Back Seat Driver's internal representation for text is preferably based on syntactic structure, not text strings. Second, the objects mentioned in the signs (cities and roads) should be entered into the discourse model to become salient for future reference. The Back Seat Driver should parse sign text by separating it into tokens delimited by commas and the word "and", and then attempt to recognize objects on the map (street names, cities, neighborhoods) from these tokens. When recognition fails, the token cannot be entered into the discourse model. When parsing fails, the spoken output will have incorrect grammar.

The Back Seat Driver does not assume that the driver will recognize the place to act (e.g. by seeing a street sign) so the driver must be told when (or where) to act. The Back Seat Driver uses timing ("Take a left here") when the driver has reached the place to act. The working prototypes calculate the place to speak by finding a distance from the intersection which is $v * (t_{speak} + t_{reaction})$, where t_{speak} is the time to speak the utterance and $t_{reaction}$ is the driver's reaction time. The time to speak depends on the number of words in the utterance. (The Dectalk synthesizer used in the prototypes speaks 180 words per minute.) Reaction time can be estimated to be two seconds.

The Back Seat driver speaks autonomously, but preferably provides means to allow it to speak on demand. The driver at any time should be able to ask for instructions immediately by, for example, pushing buttons representing "What next?" and "What now?". In addition, while following a route, a driver should be able to ask to hear the total length of the route and the remaining distance. A driver should also be able to ask to hear the name of the street the car is on and the compass direction the car is headed.

In order to generate more fluent text, the Back Seat Driver preferably keeps track of what has been mentioned. Some instructions are obvious after having been given. If the system tells the driver to go straight through a set of lights, there is no reason to repeat the instruction when actually at the lights. This is in contrast with a turn, where the driver hears advance instructions to know what to do, and immediate instructions to know when to do it. This can be important, for if the driver hears "go straight through the lights" twice, she may try to go through two sets of lights. To implement this, each instruction should be able to determine whether it is obvious after having been given once. When it is time to speak the instruction, if the instruction has already been given, and it is obvious once spoken, then it should not be spoken again.

The Back Seat Driver preferably retains a history of the route. This allows it to generate cue phrases for the instructions. If the route calls for doing the same thing twice in a row, the system uses the work "another" to indicate this doubling. This is important for polite behavior. If a passenger were to give a driver instructions by just saying "Take a right. Take a right. Take a left. Take a right.", pronouncing each the same, the passenger would be judged to be rude. The passenger's speech is not acknowledging the driver's actions or history. There are two ways for the passenger to acknowledge

the driver's work: using cue words ("Take a right. Take another right. Now take a left."), or using intonation. However, some speech synthesizers, such as the Dectalk used in the prototypes, does not support flexible control of intonation, so cue words are the only possibility.

The Back Seat Driver preferably is able to warn the driver about dangers which can be inferred from knowledge of the road network. These dangers include driving above the speed limit, driving the wrong way on a one-way street, driving too fast for an upcoming curve, driving on a one-way street that becomes two-way ahead, merging traffic, "blind" driveways ahead, speed traps, poorly repaired roads, potholes, and dangerous intersections. The Back Seat Driver preferably attempts to determine hazards by reasoning about road conditions rather than requiring them to be built in to the map database.

Lane advice includes telling the driver which lane to get into (or stay out of) when applicable. The system gives lane advice as part of the instruction when approaching an intersection where it matters. The instruction may also include advice about what lane to be in after the intersection, in preparation for the next act.

Speed advice includes warning the driver to slow down if she is travelling too fast to safely negotiate a turn. The limiting factor for angular acceleration is the driver, not the cornering ability of the car. Research has shown that the average driver will accept no more an 0.1 G radial acceleration. Radial acceleration is v^2/r where r is the turning radius of the turn. The Back Seat Driver knows the geometry of the road, so it can predict the maximum tolerable velocity for the turn. It need not tell the driver about this speed (the driver will choose a comfortable speed without being told), but it should estimate the distance required to decelerate, and tell the driver to slow down early enough to do this gently.

If the driver leaves the route, the Back Seat Driver immediately informs the driver and begins to plan a new route. Telling the driver what she did wrong prepares her for hearing new instructions, and perhaps helps her learn to better interpret the style of language that the Back Seat Driver uses.

To describe an error, the Back Seat Driver needs to look back to the last action that the driver failed to perform. It should utter a description of this action, saying e.g. "Oops, I meant for you to take a right," which does not blame the driver as in e.g. "You made a mistake. You should have taken a right." A driver might leave the route deliberately, or the error could be system's, not the drivers.

Errors will occur due to inaccuracies in the location system. The Back Seat Driver is preferably able to model the uncertainty of a position. For instance, when two roads diverge at a narrow angle, it will be unable to distinguish which was taken until some distance passes. It should probably assume that the driver made the correct choice rather than taking the risk of making a false accusation. If it reaches a place where the difference is crucial, yet unknown, it is probably better for it to confess its uncertainty, and say something like "I'm not quite sure where we are, but if you can take a right here, do it, and if not, keep going, and I'll figure things out better in a minute." Or it might ask the driver to pull over and stop (if the driver is at a place where that is safe) to allow for a better position estimate by averaging a few successive estimates.

Errors will also occur if the database is somewhat out of date. The Back Seat Driver can regain at least a little confidence by how it explains the mistake. Suppose that the Back Seat Driver intends the driver to turn onto "Apple" Street, and says "Take a right at the next light". Unbeknownst to it, a new traffic light has been installed at "Pear" Street, so the driver turns there. It is somewhat confusing if the Back Seat Driver says "I meant for you to go straight," because the driver may think that the program has not been consistent. A better message would be "I did not mean for you to turn onto Pear. I thought that the next set of lights was at Apple Street."

While the driver is following a route, the Back Seat Driver preferably adopts a persistent goal of keeping the user reassured about her progress and the system's reliability. If the Back Seat Driver were a human, this might be unnecessary, since the driver could see for herself whether the navigator was awake and attending to the road and driver. But the driver cannot see the Back Seat Driver and so needs some periodic evidence that the system is still there. One piece of evidence is the safety warnings the system gives. But if all is going well, there will not be any. Other kinds of evidence that things are going well should be provided. When the user completes an action, the Back Seat Driver can acknowledge the driver's correct action, saying something like "nice work" or "good". Also, insignificant remarks about the roads nearby, the weather and so on. can be provided. The driver then assumes that everything is going well, for otherwise the Back Seat Driver would not speak of trivial matters.

The Back Seat Driver should know about the knowledge and desires of its driver, and act differently because of this knowledge. This knowledge is preferably incorporated in a user-model.

For driver properties which do not change or change very slowly, such as colorblindness, or visual or aural acuity, it is acceptable for the Back Seat Driver to ask the user for such knowledge. However, for other driver properties, the Back Seat Driver preferably acquires a model of the user automatically, without asking or having to be told. For example, the Back Seat Driver could learn the driver's reaction time by measuring the time between its speech and the driver's operation of the controls.

The Back Seat Driver preferably learns the style of instruction giving appropriate for the driver. To learn the driver's preferences for route description requires either observation of the driver herself giving instructions or getting feedback from the driver about the instructions the system provides.

The driver can provide feedback about the suitability of the Back Seat Driver's instructions either explicitly or implicitly. One explicit indication of comprehension is how often the driver hits the "what now?" button. The system might collect long term statistics on the types of intersections where the user most often requests help, and begin to offer instructions without being asked. Just as the user can ask for more talking with the "what now" button, the Back Seat Driver should provide a "shut up" button (or other means to the same effect). The Back Seat Driver could also learn the effectiveness of its directions simply by watching the driver's performance—in particular, her errors. In this way, it can learn which cues are most effective.

Another opportunity for learning the driver's style is during the acquisition of speech recognition templates

(for user-dependent speech recognition for driver input means, described below). The new user should play the role of a "back seat driver" and give instructions, while in a car, for some route. The instructions must be given while driving either a real car or a close simulation because the form of static driving instructions is much different from real time instructions. Given some a priori knowledge about the ways that a route can be described, it is not impossible that the system could understand the instructions, and infer style from it. A difficulty here is that if the driver knows the route well, many things will seem obvious to her that would not be obvious to another person.

If the Back Seat Driver knows what the driver knows about the city, it can give better directions. A user who knows about a city no longer need instructions, she needs information about structure. The object description system preferably provides novice users a process description which emphasizes casual connections, and experts structural descriptions. Experts do not need the casual information, they can derive it for themselves.

The attributes of the user-model preferably include:

route-preference—does the driver want the fastest, shortest, or simplest route?

reassurance-period—how often should the program speak to the driver?

use-names—should the program tell the driver the names of passing streets?

congratulate-after-act—should the program make an explicity acknowledgment of correctness to the driver after each act?

obvious-to-cross-major—can the program assume that the driver will continue straight across a major intersection without being told explicitly to do so?

scoflaw—does the driver want to be warned when she is speeding?

daredevil—does the driver want warnings when driving dangerously fast?

distance-lowpass—does the driver want to be told the distance to the next action (in yards or miles, as appropriate). Most drivers do not understand distances in tenths of miles, so by default the program mentions only those distances that exceed one half mile.

The functions of the user-model preferably include:

obvious-next-segment—given a current position, is there a unique segment such that it is almost certain the driver will go there, without being told to do so?

at-major-intersection—is the current intersection one that the driver would call "major"?

extrapolate-path—try to predict the path the driver will follow in the next N seconds, assuming she does only what is obvious.

maximum-safe-speed—calculate the maximum speed at which the driver can get through an intersection. This calculation is based on finding the segment with the greatest radius of turn, and then calculating the largest speed the vehicle could have while making that turn without undergoing unacceptable sideways acceleration.

For the Back Seat Driver to decide what to say and when to say it, it preferably has a model of the vehicle performance. It must know, for example, how slowly the car should be going in order to safely make a turn. A suitably instrumented car could also measure the coefficient of friction by comparing the applied braking force and the resulting deceleration. This would allow it to adjust the time factors used in deciding when to speak.

The Back Seat Driver should understand the driver's plans and goals. When a driver enters a destination address, she is telling the system that she has the goal of getting to that address. The Back Seat Driver might guess at higher level plans from knowledge about the destination, and take actions to assist the driver with those plans. To do this, it must know what kind of thing is at the destination address. For instance, if the address provided is that of a store, the Back Seat Driver can guess that the driver is going there to purchase something, or at least to do business there. It could check the hours that the store is open.

The Back Seat Driver should help drivers to understand the route it gives. This would make the system more pleasant to use. It would also make it easier to follow routes because a driver who understands the route and the city will use that knowledge to help interpret the commands Back Seat Driver gives. A route should fit into a larger model of the city. This means that the Back Seat Driver itself must have a model of the city and should speak of the route in terms that relate it to the city. There are several opportunities to do this. At the beginning of the route, the driver might hear an overview of the route, naming the major paths followed and neighborhoods crossed. During the route, locations could be described not just as street address but in larger units of neighborhoods and districts. Orienting information can be included in instructions, or it might come between instructions, as a passing comment.

There are some additional services that the Back Seat Driver could easily provide. It should be able to give the location of a place without giving directions, it should be able to give the directions all at once, and it should be able to give directions between any two places. A driver might want to use these to tell someone else how to get to where they are.

The Back Seat Driver should be able to communicate with the outside world if the outside world is equipped to talk to it. For instance, after determining that a given parking garage is the closest or most convenient to the current destination, the Back Seat Driver could automatically phone or radio the garage and reserve a space.

The Back Seat Driver should be running on a computer embedded in the car, so that it can get more and better information about the state of the car and driver. For instance, when the next instruction is a turn, the Back Seat Driver should notice whether and when the driver turns on the turn signals. If the driver applies them too soon, it is possible (but not certain) that the driver has underestimated the distance to the turn; if applied at the "right time" then the system can take that the action has been understood; if never applied, then the driver has either misunderstood, or is driving hazardously.

The Back Seat Driver should also be integrated into the car's audio system, rather than having separate systems for voice and music. Furthermore, it should pay attention to what the driver is listening to. If the driver is listening to the radio, or playing a CD (or using a cellular telephone) the program should try to speak less often, on the grounds that the driver has implicitly indicated a preference for what to listen to. The program should suppress reminders and historical notes altogether. When it must speak, it should borrow the audio channel rather than trying to speak over it. The Back Seat Driver should also be aware of the driver's use of other controls in the car. It should defer speech

while the driver is adjusting the heat or the mirrors, for example, and suppress speaking altogether if the car makes sudden extreme changes in velocity. A driver trying to cope with an emergency situation does not need another distraction.

The discourse model preferred for the Back Seat Driver is a partial implementation of the discourse theory described by B. J. Grosz and C. L. Sidner ("Attention, intentions, and the structure of discourse" in *Computational Linguistics*, 12(3):175-204, 1986) and the theory of intonational meaning described by J. Hirschberg and J. Pierrehumbert ("The intonational structuring of discourse" in *Proceedings of the Association for Computational Linguistics*, 136-144, July 1986). Both of these articles are herein incorporated by reference. This model allows the program (or programmer) to create and manipulate discourse segments. The program specifies the contents of the discourse segment (both the syntactic form and the list of objects referenced) and the implementation of the model does the following: generates prosodic features to convey discourse structure; inserts discourse segment into intentional structure; and maintains attentional structure—adding new objects when mentioned and removing old objects when replaced. In addition it includes some useful low-level tools for natural language generation: search of attentional structure for occurence of co-referential objects; conjugation of verbs; generation of contracted forms; and, combination of sentences as "justification" sentences (e.g. "get in the right lane because you are going to take a right.") and sequential sentences ("Go three blocks, then turn left"). In order to use the discourse package the programmer must explicitly declare all semantic types used by the program, so in this case there are declarations for all objects which pertain to driving, such as street names, bridges, rotaries, stop lights and so on.

SPEECH GENERATOR

In the working prototypes of the Back Seat Driver, speech generation is performed by Dectalk, a commercial text-to-speech speech synthesizer, which is cabled to the main computing apparatus.

An alternative to synthesized speech is digitized speech, which is easier to understand than synthetic speech. Digitized speech, however, requires a great deal of storage space. There are more than 2000 different street names in Boston. Allowing another 500 words for the actual instructions, and assuming an average duration of 0.5 seconds for each word, coding this vocabulary at 64 kilobits per second would require 10 megabytes of speech storage. Given a Back Seat Driver that uses a CD-ROM for the map, the digitized speech could be stored on the disk as well. Coded speech would be more intelligible than synthesized speech, and less costly, but not as flexible. For, example, it would be impossible to read electronic mail using only stored vocabulary speech.

The generated speech is spoken to the driver through some kind of speaker system in the car. In a simple embodiment, the speaker system of the car radio is used.

DRIVER INPUT MEANS

Means for the driver to communicate with the backseat driver are required. For example, the driver must be able to enter destination addresses, request instructions or a repeat of instruction, and inform the Back Seat driver when an instruction cannot be carried out for some reason. In embodiments where the computing apparatus is installed in the automobile, a computer keyboard can be adapted to provide this communication means.

In one working prototype of the Back Seat Driver, the computing apparatus is not installed in the automobile, but is accessed through a cellular telephone. In this embodiment, the driver communicates with the Back Seat Driver by using the cellular telephone keypad. Address entry is achieved by first entering the digits, then a number sign, then spelling the street name using the letters on the telephone keypad. The letters "Q" and "Z" are on the "6" and "9" keys, respectively, and the space character is on "1", which is otherwise unused. These keys are sufficient to spell any street name in Boston. The spelling rules can be easily expanded to enter street names with other characters in them, for example, "4th Street".

In the implementation, spelling a street name requires only one button push for each letter, even though there are three letters on each key. This is because of the redundancy in street names, which are pronounceable words, not arbitrary strings. There are 37 pairs of street names in Boston with the same "spelling" in the reduced "alphabet". An example is "Flint" and "Eliot", both encoded as "35468". This is only one percent of the 2628 names of streets in Boston, so collisions are rare. This technique appears workable even for larger sets of names. When the entire word list of the Brown corpus is encoded, there are still only 1095 collisions in the 19,837 words (5.5%).

If a name collision occurs, the Back Seat Driver reads the list of possibilities, and asks the driver which one was meant. This is very rare. A more common problem is that street names are duplicated. When this happens, the Back Seat Driver asks the user a series of questions to reduce the list to a single choice. It tries to ask the fewest questions possible. It asks the user to choose from a list of street types, if that is sufficent to resolve the question, and otherwise from a list of the containing cities (or neighborhoods, if there are two instances within a single city). To select from a list, the Back Seat Driver reads the contents, asking the user to push a button when the desired choice is read.

The Back Seat Driver would be much easier to use if the driver could simply talk to it instead of using a keyboard or keypad. Speech recognition could be used for driver input means, however, address entry is a difficult task for speech recognition for the same reason it is hard for a human to understand machine speech—there are few constraints on a name. No speech recognizer available today can handle a 3000 word vocabulary with acceptable error rates. The car would also have to be stopped while the driver was speaking, because noise in moving cars for frequencies below 400 Hz can exceed 80 dB.

Back Seat Driver could also use speech recognition to replace the "What now?" and "What next?" buttons. This is a more tolerant application for speech recognition because there are fewer words to recognize.

SYSTEM PROCESSES

The Back Seat Driver carries out three separate tasks, each of which is executed by its own process. All processes share the same address space, so all variables and functions are accessible in every process, and no special mechanism for interprocedure call is required. Where necessary for synchronization, Back Seat Driver uses

queues or locks. All three processes are simple, infinite loops. The system processes are illustrated in FIG. 2.

The user process is the main process of the Back Seat Driver. It is this process which finds routes and gives instructions to the driver. The user process manages a list of goals. Each time around the loop, it selects a goal to work on, and does something to achieve the goal, if possible. The user process is connected to the speech generator, since that is how it talks to the driver.

The navigator process maintains an estimate of the current position and velocity of the car. It is connected to the position sensor by a serial line. Preferably, packets arrive from the position sensor several times a second. The navigator converts the data in the packets from the position sensor format to the format used by the Back Seat Driver.

There are two minor processes which assist the navigator process: The average speed process computes the running average speed of the vehicle over the last five seconds. It could be made part of the navigator process, but is distinct because it is more convenient that way. The position sensor monitor process keeps track of how often packets arrive. If packets do not arrive when scheduled, it should set a flag to indicate this to inform the driver if the position sensor ceases to work properly.

The control process is responsible for controlling the Back Seat Driver as a whole. The control process is connected to driver input means for entering, for example, the destination and requesting additional instructions while driving (e.g. the "What now"?, "What next?" and "I can't do it" features.) Other functions of the control process are useful in a research prototype, but should not be required in a commercial embodiment of the Back Seat Driver. One such function is debugging.

The user process is a goal-driven perpetual loop which seeks to use the resources available to it to satisfy as many goals as possible as quickly as possible, devoting resources first to those goals which are of greatest importance. There are two aspects to this process, goal structures (which names goals) and goal-functions (which tell how to accomplish them). A goal is just a name, a priority (a number), and a set of slots (parameters). Thus for instance a typical goal would be (GET-TO-PLACE<140 Elm Street>), where the goal has one slot, namely the destination. A goal-function is a function which is possibly able to achieve a goal. When a new type of goal is defined, the programmer also tells the system which goal functions can possibly meet it, and later, when the system tries to accomplish a goal it selects from this list.

The goal loop is a three step process. 1) Check to see whether there are any newly added goals. The driver can add a goal by hitting a key, and the system can also add goals. 2) Find the most important goal to work on. 3) Work on that goal. In general, systems should use resources in the most efficient manner possible. For the Back Seat Driver, the only resource is speaking time. The only way the Back Seat Driver can accomplish any of its goals is by speaking. Speech is a resource because the program can only say one thing at a time, and speaking takes a finite time. It is also important to note that spoken utterance has a useful effect only if completely spoken, so it is not helpful to begin to speak if there is not time to complete the speech.

The protocol for a goal function preferably includes the following:

progressable?—Is the goal able to accomplish anything at this time?

resource-used—If it runs now, what resources will it want to use?

maximum-time-of-resource—If it runs now, how long (in seconds) will it need each resource?

minimum-time-to-resource—The minimum time that it can estimate until it may again need this resource, and the priority of its use at that time.

In the working prototypes of the Back Seat Driver, the list of all goals is stored in the global variable *goals*. The goal loop and goal structures are defined in the file goals.lisp. The various goals and goal functions of the Back Seat Driver are defined in the files main.lisp, route-goals.lisp, and get-to-place.lisp. All goals which use speech are built from the speech-goal object defined in speech-goal.lisp. The speech resource itself is defined in speech-resource.lisp.

The goal or function which gets a user to a destination is called GET-TO-PLACE. An explanation of this goal will illustrate the goal mechanism in more detail, as well as illustrate how this most important function of Back Seat Driver is implemented. The goal GET-TO-PLACE, has two slots, destination which is the location the user wants to get to, and route which is the route the Back Seat Driver intends to use to get there.

The driver adds the goal to the system goal list by striking a key. When the goal is first created, the destination is not known (the destination slot is empty), so the goal function for GET-TO-PLACE creates a sub-goal, GET-DESTINATION, and adds it to the goal list. Now there are two goals on the goal list, GET-TO-PLACE and GET-DESTINATION, but only the second is progressable, because when a goal has a sub-goal, it is not allowed to run until the sub-goal finishes. Therefore, the only progressable goal is GET-DESTINATION, which attempts to get a destination by asking the user to enter an address. If the user fails to do so, the subgoal fails, which in turn causes GET-TO-PLACE to fail, and the Back Seat Driver says "Travel cancelled". Otherwise, it writes the destination into the destination slot of the GET-TO-PLACE goal. Now that the sub-goal is complete, GET-TO-PLACE can once again make progress. This time it finds that the route slot is empty, and again calls for the sub-goal GET-ROUTE, which calculates a route. When this is complete a third subgoal is invoked, namely FOLLOW-ROUTE.

The goal function for FOLLOW-ROUTE gets the driver to the destination by speaking instructions. If something goes wrong (for example if the driver makes a mistake) then the subgoal fails. But this does not make GET-TO-PLACE give up. Instead, it erases the route slot, and simply finds a new route, and then tries FOLLOW-ROUTE again. This continues, no matter how many times things go astray, until either FOLLOW-ROUTE succeeds, or the driver cancels the trip.

The goal FIND-SERVICE is similar to GET-TO-PLACE except the driver selects a kind of service (for example, a gas station), and then the Back Seat Driver finds the closest provider of that service, and then finds a route to it. Following that route is done by FOLLOW-ROUTE in the same way as for GET-TO-PLACE.

The FOLLOW-ROUTE goal function gets the user to her destination by giving spoken instructions. There are several reasons it might speak:

at the beginning, to alert the driver

to give an instruction in advance, so the driver will be ready

to give an instruction when it is time to do it

to confirm that the driver has correctly carried out an instruction

to inform the driver of her arrival at the destination

to reassure the driver that she is still on route

to inform the driver of a mistake

to warn the driver that she is driving so fast that the program cannot keep up.

FOLLOW-ROUTE decides the next reason for speaking by first trying to locate the current position on the path. If the position is not on the path (more precisely, if the current segment does not occur anywhere on the path) then the driver has left the path (or the position sensor has made an error). Otherwise, FOLLOW-ROUTE determines what instruction must be next executed by calling the function next-driver-instruction.

The goal function protocol requires that FOLLOW-ROUTE support the goal function minimum-time-to-resource, which estimates the minimum time until FOLLOW-ROUTE will next speak. This time depends upon the reason for the next speaking. FOLLOW-ROUTE speaks immediately when beginning, confirming, warning, or finishing the route. When the driver is off the route, FOLLOW-ROUTE waits a few seconds before speaking, just in case the driver's departure from the route is actually a temporary error by the position sensor.

Given that the driver is on the path, FOLLOW-ROUTE determines when to speak by calculating the position where it must begin speaking the instruction text, then estimating the time required to reach that position at the driver's current speed. As the driver's speed changes, so will this estimated time. The position to begin speaking is calculated by first finding the position where the instruction is executed, then moving back a distance to allow the Back Seat Driver time to speak the text and the driver to react to it.

The Back Seat Driver can also give instructions in advance, if desired. It does this in much the same way, except that it adds an additional number of seconds (normally thirty) to the time estimate, and so begins to speak much sooner. When it gives instructions in advance the instruction text is longer because the program has more time to speak.

When the driver leaves the route FOLLOW-ROUTE starts a timer. If the driver has not returned to the route by the time the timer goes off (at present, two seconds) then FOLLOW-ROUTE checks for a possible mistake. In describing the mistake, it attempts to characterize what the driver actually did as well as what the program intended the driver to do. It is able to do this because in the main loop it stored the last position that the driver was on when last on the route.

Goals may interrupt lower priority goals by requesting the speech resource to interrupt the lower priority goal. Interruption stops the speech synthesizer immediately. The interrupted goal is informed of the interruption, and can react as it chooses. There is no way for the goal to know whether any of its words were actually spoken, so it has to start all over. Most goals attempt to run again as soon as possible. The assumption is that the interruption occurred because the user started some higher priority goal after learning how to do so through the help command.

The system treats "repeat the last statement" as a goal, rather than as a special purpose function, except that the importance of this goal is set to the value of the last goal spoken (the goal whose utterance is being repeated). This guarantees that if some more important goal desires to speak, it will be able to. A repetition of an utterance is no more important than it was originally.

Goals can be temporary or presistent. Temporary goals can be satisified, but persistent goals never can be. All system initiated goals are persistent. The system goals include warning the driver of dangers ahead (WARN-DRIVER) and informing the user of new electronic mail or other messages (if the computer apparatus of the Back Seat Driver is connected to the outside world). These goals can never be satisified. The driver's safety should always be preserved and mail or messages can arrive at any time.

CELLULAR PHONE EMBODIMENT

The Back Seat Driver is preferably an in-car navigation system, but in some embodiments, it may be desirable to not have the entire computing apparatus installed in the car. This is the case if the computing apparatus is too large or if a number of cars are to share a single computing apparatus.

For such embodiments, two cellular phones installed in the car can be used to transmit data from the car to the computing apparatus, and to receive voice from the speech generator in the computing apparatus. In this embodiment, data from the position sensor installed in the automobile is sent through a cellular phone in the car equipped with a modem to a phone connected to the computing apparatus via a modem. The voice generating apparatus of the computing apparatus sends speech over another phone to a second cellular phone installed in the automobile.

This embodiment has been implemented in a working prototype, using a large workstation computer (a Symbolics Lisp Machine). In this implementation, a position sensor installed in the car estimates vehicle position, heading, and velocity, and sends a data packet, once per second, through a modem to the workstation. The workstation sends characters to a Dectalk speech synthesizer, which in turn sends voice over a second phone to the driver.

Nearly everyone who has used a cellular phone knows how noisy they are. Cross talk is common and noise bursts and signal loss make it hard to hear. A sufficiently bad noise burst can even cause the cellular system to terminate the call. The problems for data transmission are even worse. By its very nature, cellular radio transmission is subject to multi-path interference, which causes periodic fades as the antenna moves in and out of anti-nodes. In addition to this fading, there is a complete loss of audio signal for as long as 0.9 seconds when the phone switches from one cell site to another (hand off).

An attempt to use an ordinary (land-line) modem from the car was unsuccessful. In the prototype, a Worldlink 1200 from Touchbase Systems was used in the car, with a Morrison and Dempsey AB1 data adapter, and an NEC P9100 phone, boosted to 3 watts. At the base station, both a Practical Peripherals 2400 and a Hayes Smartmodel 1200 were used. Even at 300 baud the connection was too noisy to use. Worse, connections seldom lasted more than five minutes. In all cases, the "loss of carrier" register (S10) was set to its maximum value, 20 seconds. Loss of carrier signal alone

cannot explain why the connections dropped. The modems were capable of tolerating a complete loss of audio for up to twenty seconds.

Better results were found using an error correcting modem (The "Bridge") made by the Spectrum Cellular Corporation. This modem uses a proprietary protocol (SPCL) for error correction. The Spectrum product virtually eliminated noise, at the price of a lower data transmission rate. According to the protocol, the transmitting modem groups characters into packets that include error correction bytes. If only a few errors occur, the receiving modem repairs the data and acknowledge receipt. If there are many errors, the packet is retransmitted. If the sending modem has to retransmit too often it makes the packets smaller, on the assumption that a smaller packet has a better chance of success. This is less efficient, since packets have a fixed overhead, the percent of the channel used by data decreases. When conditions improve the modem increases packet size again. In theory, the modem can transmit at 120 characters per second, but tests made by recording the time required to receive the three characters of an odometer sequence demonstrated that the average value is closer to 30 characters per second. This sequence is transmitted once per second. The mean for durations for the three character sequences is 94 milliseconds, which is 31 milliseconds per character, or 32 characters per second.

Even with the cellular modem, calls are sometimes dropped. The call durations are usually long enough for a successful trip with the Back Seat Driver. Voice calls are dropped at about the same rate as data calls.

The protocol used by the Spectrum modem acknowledges all data transmitted. If the acknowledgment is not received, it retransmits the data until acknowledged. Under adverse conditions, this can result in an arbitrarily long delay. This is a problem when real-time data is transmitted. Observation of the Back Seat Driver shows that sometimes the system will "freeze" for from one to ten seconds. During this time, the car of course continues to move. If these freezes occur near decision points, the driver may go past the intersection without being told what to do. At 20 miles per hours a car travels nearly 45 meters in five seconds. The navigation system in the car sends a packet once every second. Most packets arrive within a second, but a few are delayed, some by up to ten seconds. (These delays may also arise from delays at the workstation. Lisp Machines are not noted for real-time response.)

It would be better to have a protocol which guarantees to deliver data intact and free of errors, if it delivers it at all, but does not guarantee to deliver the data. Real time data is only valuable in real time, and time spent retransmitting old data is taken away from ever, more valuable data. Such a protocol modification is feasible for the Spectrum product.

What is claimed is:

1. An automobile navigation system which produces spoken instructions to direct a driver of an automobile to a destination in real time comprising:

computing apparatus for running and coordinating system processes,

driver input means functionally connected to said computing apparatus for entering data into said computing apparatus, said data including a desired destination,

a map database functionally connected to said computing apparatus which distinguishes between physical and legal connectivity,

position sensing apparatus installed in the automobile and functionally connected to said computing apparatus for providing said computing apparatus data for determining the automobile's current position,

a location system functionally connected to said computing apparatus for accepting data from said position sensing apparatus, for consulting said map database, and for determining the automobile's current position relative to the map database,

a route-finder functionally connected to said computing apparatus, for accepting the desired destination from said driver input means and the current position from said location system, for consulting said map database, and for computing a route to the destination,

a discourse generator functionally connected to said computing apparatus for accepting the current position from said location system and the route from said route finder, for consulting said map database, and for composing discourse including instructions and other messages for directing the driver to the destination from the current position,

a speech generator functionally connected to said discourse generator for generating speech from said discourse provided by said discourse generator, and

voice apparatus functionally connected to said speech generator for communicating said speech provided by said speech generator to said driver.

2. The automobile navigation system of claim 1 wherein said map database comprises a set of straight line segments and a set of nodes, each endpoint of each segment being a pointer to a node representing the coordinates of the endpoint and the set of other segments which are physically and legally connected to that endpoint.

3. The automobile navigation system of claim 1 wherein said map database is based on DIME files of the United States Census extended to represent physical and legal connectivity.

4. The automobile navigation system of claim 3 wherein said DIME file is further extended to distinguish bridges, underpasses, tunnels, rotaries, and access ramps from other street types.

5. The automobile navigation system of claim 1 wherein said map database is based on TIGER files of the United States Census and United States Geological Survey extended to represent physical and legal connectivity.

6. The automobile navigation system of claim 5 wherein said TIGER file is further extended to distinguish bridges, underpasses, tunnels, rotaries, and access ramps, from other street types.

7. The automobile navigation system of claim 1 wherein said map database comprises a three-dimensional representation of street topology.

8. The automobile navigation system of claim 1 wherein said map database includes measures of street quality.

9. The automobile navigation system of claim 1 wherein said map database distinguishes divided streets.

10. The automobile navigation system of claim 1 wherein said map database includes landmarks such as signs, traffic lights, stop signs and buildings.

11. The automobile navigation system of claim 1 wherein said map database includes lane information.

12. The automobile navigation system of claim 1 wherein said map database includes speed limits.

13. The automobile navigation system of claim 1 wherein said map database includes expected rate of travel.

14. The automobile navigation system of claim 1 wherein said map database includes time-dependent legal connectivity.

15. The automobile navigation system of claim 1 wherein said map database includes turn difficulty.

16. The automobile navigation system of claim 1 wherein said map database includes vehicle street, lane, and height restrictions.

17. The automobile navigation system of claim 1 wherein said map database includes traffic light cycles.

18. The automobile navigation system of claim 1 wherein said map database distinguishes where right turn on red is allowed.

19. The automobile navigation system of claim 1 wherein said map database includes a database of service locations.

20. The automobile navigation system of claim 1 wherein said map database includes a listing of famous places by name.

21. The automobile navigation system of claim 1 further comprising means for updating said map database.

22. The automobile navigation system of claim 1 further comprising means for updating said map database by radio broadcast.

23. The automobile navigation system of claim 1 wherein the map has minimum accuracy of 10 meters.

24. The automobile navigation system of claim 1 wherein said route finder is based on a best-first search algorithm.

25. The automobile navigation system of claim 1 wherein said route finder is based on an A* algorithm.

26. The automobile navigation system of claim 1 wherein said route finder is based on an A* algorithm modified to find a route in less time.

27. The automobile navigation system of claim 1 wherein said route finder is adapted to find a best route according to any one of three cost metrics: distance, speed, simplicity.

28. The automobile navigation system of claim 1 wherein said route finder is adapted to calculate a new route if the driver or vehicle navigation system makes an error or if the route is unnavigable due to unforeseen circumstances, wherein said new route does not simply backtrack to the point of the error if a better route from the current location exists.

29. The automobile navigation system of claim 1 wherein said route finder is adapted to calculate a new route while the automobile is in motion, wherein said new route will begin from the location of the automobile at the time the calculation of the new route is completed.

30. The automobile navigation system of claim 29 wherein an estimated time to find a new route is multiplied by the velocity of the automobile to calculate the position from which the new route should start.

31. The automobile navigation system of claim 30 wherein said estimated time to find a new route is calculated by multiplying the distance between the starting and ending points of the new route by a constant.

32. The automobile navigation system of claim 1 wherein said location system is a position-keeping (dead-reckoning) system.

33. The automobile navigation system of claim 1 wherein said location system is a hybrid of position-keeping and position-finding systems.

34. The automobile navigation system of claim 1 wherein said location system employs map matching.

35. The automobile navigation system of claim 1 wherein said position sensing apparatus comprises displacement and direction sensors installed in the automobile.

36. The automobile navigation system of claim 1 wherein said position sensing apparatus measures displacement with an odometer.

37. The automobile navigation system of claim 1 wherein said position sensing apparatus measures direction with a magnetic compass.

38. The automobile navigation system of claim 1 wherein said position sensing apparatus measures direction by monitoring the turning of the steering wheel.

39. The automobile navigation system of claim 1 wherein said position sensing apparatus measures direction with a differential odometer.

40. The automobile navigation system of claim 1 wherein said position sensing apparatus measures direction with a gyroscope.

41. The automobile navigation system of claim 1 wherein said discourse generator is based on an object-oriented programming methodology.

42. The automobile navigation system of claim 1 wherein each intersection in a route is classified into one type in a taxonomy of intersection types, and the disclosure generated in relation to each said intersection depends on its type.

43. The automobile navigation system of claim 42 wherein said taxonomy of intersection types includes continue, forced-turn, U-turn, enter, exit, onto-rotary, stay-on-rotary, exit-rotary, fork, turn, and stop.

44. The automobile navigation system of claim 42 wherein said discourse generated further depends on a description function for each intersection type which generates a description given the length and tense of the desired description and the position along the route from which an instruction is to be given.

45. The automobile navigation system of claim 1 wherein said discourse generated comprises a long description of an act given substantially before the act is to be performed and a short description given at the time the act is to be performed.

46. The automobile navigation system of claim 45 wherein said long descriptions includes cues.

47. The automobile navigation system of claim 46 wherein said cue is a landmark.

48. The automobile navigation system of claim 1 wherein said driver input means includes means for said driver to demand immediate instructions, or clarification or repetition of instructions already provided.

49. The automobile navigation system of claim 1 wherein said driver input means includes means for said driver to indicate to said automobile navigation system that a given instruction provided by said system is impossible to complete for some reason and that a new route must be calculated.

50. The automobile navigation system of claim 1 wherein said driver input means comprises a voice recognition system to allow at least some driver input to be spoken.

51. The automobile navigation system of claim 1 wherein said automobile navigation system records a history of the route and the discourse already generated and uses this knowledge to generate cues for future discourse and make future discourse more understandable.

52. The automobile navigation system of claim 1 wherein said automobile navigation system warns drivers of dangers inferred from knowledge of the road network.

53. The automobile navigation system of claim 1 wherein said automobile navigation system informs a driver if an error has been made as detected by the location system.

54. The automobile navigation system of claim 1 wherein said discourse generator is responsive to a user-model stored in said computing apparatus to customize discourse to the requirements and preferences of said driver.

55. The automobile navigation system of claim 1 wherein said speech generator is a speech synthesizer.

56. The automobile navigation system of claim 1 wherein said speech generator uses digitized speech.

57. The automobile navigation system of claim 1 wherein said computing apparatus is not installed in the automobile, and said automobile navigation system further comprises means for communication between said computing apparatus and the automobile navigation system components installed in the automobile.

58. The automobile navigation system of claim 57 wherein said means for communication is two cellular phones in said automobile, one of which is connected to a modem, and two phones connected to said computing apparatus, one of which is connected to a modem, whereby one data channel and one voice channel between said automobile and said computing apparatus is created.

* * * * *

EXHIBIT

2

Excerpts from the Doctoral Thesis of James Raymond Davis

This document was produced in its entirety during both the Claim Construction (Ex. I) and Inequitable Conduct (Ex. 2) briefing. To avoid unnecessary duplication, Harman will provide a complete copy of this document at the Court's request.

IW 1415424

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

February 07, 2006

THIS IS TO CERTIFY THAT ANNEXED IS A TRUE COPY FROM THE RECORDS OF THIS OFFICE OF THE FILE WRAPPER AND CONTENTS OF:

APPLICATION NUMBER: ***07/565,274***
FILING DATE: ***August 09, 1990***
PATENT NUMBER: ***5,177,685***
ISSUE DATE: ***January 05, 1993***

By Authority of the
Under Secretary of Commerce for Intellectual Property
and Director of the United States Patent and Trademark Office

N. Williams
N. WILLIAMS
Certifying Officer

PART (2) OF (3) PART(S)

0001

5177685

07/565274 PATENT DATE JAN 05 1993 PATENT NUMBER

SERIAL NUMBER	FILING DATE	CLASS	SUBCLASS	GROUP ART UNIT	EXAMINER
07/565,274	08/09/90	364		2	

JAMES R. DAVIS, NORTH CAMBRIDGE, MA; CHRISTOPHER M. SCHMANDT, MILTON, MA.

CONTINUING DATA*******************
VERIFIED
NONE

FOREIGN/PCT APPLICATIONS**********
VERIFIED
NONE

CERTIFICATE
MAR 1 5 1994
OF CORRECTION

FOREIGN FILING LICENSE GRANTED 09/22/90

Foreign priority claimed / 35 USC 119 conditions met	AS FILED	STATE OR COUNTRY	SHEETS DRWGS	TOTAL CLAIMS	INDEP CLAIMS	FILING FEE RECEIVED	ATTORNEY'S DOCKET NO.
☐ yes ☒ no / ☐ yes ☒ no; Verified and Acknowledged, Examiner's Initials	→	MA	2	58	1	$ 946.00	

SAM PASTERNACK
CHOATE, HALL & STEWART
EXCHANGE PLACE
53 STATE STREET
BOSTON, MA 02109

AUTOMOBILE NAVIGATION SYSTEM Using Real Time Spoken Driving Instructions

U.S. DEPT. OF COMM. - PAT. & TM. OFF. - PTO-436L

PARTS OF APPLICATION FILED SEPARATELY

NOTICE OF ALLOWANCE MAILED

PREPARED FOR ISSUE

CLAIMS ALLOWED

Total Claims 58

ISSUE FEE

DRAWING

Sheets Drwg. 3

Primary Examiner

ISSUE CLASSIFICATION

Class 364 Subclass 443

ISSUE BATCH NUMBER

Label Area

WARNING: The information disclosed herein may be restricted. Unauthorized disclosure may be prohibited by the United States Code Title 35, Sections 122, 181 and 368. Possession outside the U.S. Patent & Trademark Office is restricted to authorized employees and contractors only.

0002

9

Back Seat Driver: voice assisted automobile navigation

by

James Raymond Davis

B.S.A.D., Massachusetts Institute of Technology (1977)

Submitted to the Media Arts and Sciences Section
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1989

Signature of Author ....
Media Arts and Sciences Section
August 4, 1989

Certified by ....
Nicholas P. Negroponte
Professor of Media Technology
Thesis Supervisor

Accepted by ....
Stephen A. Benton
Chairman, Departmental Committee on Graduate Students

Back Seat Driver: voice assisted automobile navigation

by

James Raymond Davis

B.S.A.D., Massachusetts Institute of Technology (1977)

Submitted to the Media Arts and Sciences Section
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1989

Signature of Author ....
Media Arts and Sciences Section
August 4, 1989

Certified by ....
Nicholas P. Negroponte
Professor of Media Technology
Thesis Supervisor

Accepted by ....
Stephen A. Benton
Chairman, Departmental Committee on Graduate Students

Back Seat Driver: voice assisted automobile navigation

by

James Raymond Davis

Submitted to the Media Arts and Sciences Section
on August 4, 1989, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The Back Seat Driver is a computer navigation assistant for drivers in a city. It differs from earlier navigation programs by using speech, rather than graphics, to give instructions. The advantages of speech are that the driver's eyes are left free for driving and that the spoken directions contain information not easily portrayed in pictures. The program talks about the features of the road in the same way the driver sees them, giving the impression that the program is actually is in the car.

Driving instructions are modeled after those given by people. The two issues for spoken directions are *what to say* (content) and *when to say it* (timing). The content of the instructions tells the driver what to do and where to do it. The program has a large taxonomy of intersection types, and chooses verbs to indicate the kind of intersection and the way of moving through it. The instructions refer to landmarks and timing to tell the driver when to act.

Timing is critical because speech is transient. Drivers hear instructions just in time to take the required action, and thus need not remember the instruction or exert effort looking for the place to act. The program also gives instructions in advance, if time allows, and the driver may request additional instructions at any time. If the driver makes a mistake the program describes the mistake, without casting blame, then finds a new route from the current location.

Street map data bases for navigation programs must distinguish between *physical* connectivity (how pieces of pavement connect) and *legal* connectivity (whether one can legally drive onto a physically connected piece of pavement). Legal connectivity is essential for route finding, and physical connectivity for describing the route. The database must also contain all landmark information, since the program has no "eyes".

The Back Seat Driver is an actual working prototype. It has successfully guided drivers unfamiliar with Cambridge to their destinations. Although much work remains, it is easy to foresee a practical implementation in the future.

Thesis Supervisor: Nicholas P. Negroponte
Title: Professor of Media Technology

EXHIBIT

2 (part 2)

Excerpts from the Doctoral Thesis of James Raymond Davis

This document was produced in its entirety during both the Claim Construction (Ex. I) and Inequitable Conduct (Ex. 2) briefing. To avoid unnecessary duplication, Harman will provide a complete copy of this document at the Court's request.

Acknowledgments

I owe gratitude to many people for help with this work. In particular I would like to thank: Phil Rittmueller and NEC Home Electronics (USA) Inc., who sponsored it; Symbolics Inc., Digital Equipment Corporation, the Nippon Telephone and Telegraph company, and DARPA, which sponsored my earlier graduate studies; Brewster Kahle, who provided the original inspiration for my work on computerized direction giving in the summer of 1985; Thinking Machine Corporation, where the work was conducted; and Tom Trobaugh, who began that study with me. I thank those who provided me with helpful information: James A. O'Connell, Jr. and Peter Naghavi of the Somerville Department of Traffic and Parking, Jane Kent of the Cambridge Department of Traffic, and George Hawat of the Boston Transportation Department, for providing traffic light data bases; Doris Walter of the Massachusetts District Commission for maps of rotaries; Doug Milliken for explanations of car dynamics; Don Cooke of Geographic Data Technology and Joel Sobel of the Census Bureau for information on the DIME and TIGER map formats; David Pietraszewski of the United States Coast Guard and Michael Fisher of Trimble Navigation for information about GPS. The photo on page 60 was taken by Kyle G. Peltonen, and is reprinted by permission of The Tech, the Mr. Boffo cartoon is reprinted courtesy of the Tribune Media Services. I thank also Weng-yew Ko for building hardware interfaces; Elaine McNair at NEC for removing many obstacles; Michael Sullivan of Spectrum Communication for solving communications problems; Greg Parro of Cellular Phone Services for excellent service with the tele-

phones; Hidehiro Matsumoto of NEC for translations to and from the Japanese; Janette Noss and Elaine McCarthy, who keep the Architecture Machine Group running; Anna Korteweg, Walter Bender, Gary Drescher, and Carol Strohecker for careful editing, and Shawn T. Williams and Greg Grove, who helped by participating in UROP work. I wish to thank all those who taught me of computational linguistics, especially Barbara Grosz, Julia Hirschberg, and Janet Pierrehumbert. They of course bear no responsibility for my failure to fully grasp the material they attempted to teach me. My committee members were Steven Benton, Mike Lesk, Nicholas Negroponte, and Chris Schmandt. None of these people, of course, are responsible in any way for the ways I have misunderstood their ideas or abused their work, or for my failure to heed their advice. Finally I must specially mention Nicholas Negroponte, who created the Media Lab where this work was possible, and Chris Schmandt, the director of the Speech Research Group, who has been directly involved in everything I have done here at the Media Lab. A full list of his contributions to my studies would require many pages.

Those I have forgotten or otherwise neglected, I can only ask to deepen my debt by forgiving my shoddy memory and inexpressive words.

This thesis is dedicated to my son, Adam, may he find his way back soon.

EXHIBIT
3

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF MASSACHUSETTS

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, **Plaintiff,** **v.** **HARMAN INTERNATIONAL INDUSTRIES, INCORPORATED,** **Defendant.**	**Case No: 05-10990 DPW** **Hon. Douglas P. Woodlock**

MIT'S SUPPLEMENTAL RESPONSES TO
HARMAN'S INTERROGATORY NOS. 1-5, 8, 11, AND 13-14

Pursuant to Rules 26 and 33 of the Federal Rules of Civil Procedure, Plaintiff, Massachusetts Institute of Technology ("MIT") submits the following supplemental responses and objections to Harman International Industries, Incorporated's ("Harman's") Interrogatory Nos. 1-5, 8, 11, 13, and 14 (the "Interrogatories").

GENERAL OBJECTIONS

MIT herein incorporates by reference its General Objections as set forth in MIT's First Supplemental Response to Harman's Interrogatory Nos. 11-15.

SPECIFIC OBJECTIONS AND RESPONSES

INTERROGATORY NO. 1

State MIT's proposed construction, and all bases supporting such construction, of the following element of claim 1 of U.S. Patent No. 5,177,685 (the "Patent-In-Suit"): "a map database functionally connected to said computing apparatus which distinguishes between physical and legal connectivity."

1 year before the filing date of U.S. Patent No. 5,177,685. Mr. Davis' thesis (*see* HAR 001479) also notes that the Back Seat Driver had been used more than 1 year before the filing date of U.S. Patent No. 5,177,685. For each asserted claim of United States Patent No. 5,177,685, identify each and every limitation of the claim that MIT contends was not embodied in a field trial prior to August 9, 1989, and explain in detail all bases for any contention by MIT that such field trials do not render each asserted claim of the '685 patent invalid under 35 U.S.C. § 102(b).

SECOND SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 11

MIT objects to this Interrogatory as overly broad, unduly burdensome, and not reasonably calculated to lead to the discovery of admissible evidence. MIT further objects to this Interrogatory because it calls for a legal conclusion with respect to validity. MIT further objects to this Interrogatory because it seeks information protected by the attorney-client privilege, work product doctrine, and/or other applicable privileges or immunities.

Subject to and without waiving the foregoing general and specific objections, MIT incorporates by reference its Supplemental Response to Interrogatory No. 11 as if fully set forth herein.

MIT further identifies pages 1-138 of MIT's 30(b)(6) deposition testimony as responsive to this Interrogatory.

INTERROGATORY NO. 13

For each claim of U.S. Patent No. 5,177,685, identify the date(s) on which the subject matter recited therein was first completely conceived, and identify by Bates number all documents or other material that evidence all such date(s) in any way.

THIRD SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 13

MIT objects to this Interrogatory as overly broad, unduly burdensome, and not reasonably calculated to lead to the discovery of admissible evidence. MIT further objects to this Interrogatory as premature to the extent that it calls for a legal conclusion with respect to conception. MIT further objects to this Interrogatory to the extent that it mischaracterizes the legal standard for conception. MIT further objects to this Interrogatory because it seeks information protected by the attorney-client privilege, work product doctrine, and/or other applicable privileges or immunities.

Subject to and without waiving the foregoing general and specific objections, MIT objects to the phrase "completely conceived", but states that the subject matter of the '685 patent was conceived before the filing date of the application on which the '685 patent issued. The details of the conception were fully described in answer to numerous questions to the inventors propounded during the deposition testimony of Dr. James R. Davis, Ph.D, Christopher M. Schmandt and MIT pursuant to Rule 30(b)(6) of the Federal Rules of Civil Procedure and in response to this Interrogatory, and those answers are herein incorporated by reference.

MIT further states that claims 1-4, 7-10, 14, 16, 19, 24, 27, 28, 42-44, 48, 55, and 57-58 were conceived at least as early as April 1988. MIT further states that claims 5, 6, 11-13, 15, 17, 18, 20-23, 25, 26, 29-41, 45-47, 49-54, and 56 were conceived at least as early as June of 1989.

MIT identifies documents bearing Bates numbers MIT00433-MIT00947, MIT01101-MIT01102, MIT01370-MIT01378, MIT01955-MIT02002, and MIT02155-MIT02274 as responsive to this Interrogatory. MIT further identifies pages 83-90, 104, 145-162, 179, and 287 of Mr. Schmandt's deposition transcript as responsive to this Interrogatory. MIT further identifies pages 79-91, 129, 167-169, 205, and 219-238 of Dr. Davis' deposition transcript as

responsive to this Interrogatory. MIT further identifies pages 1-138 of MIT's 30(b)(6) deposition testimony as responsive to this Interrogatory.

INTERROGATORY NO. 14

For each claim of U.S. Patent No. 5,177,685, identify the earliest date(s), if any, on which the subject matter recited therein was first actually reduced to practice, and identify by Bates number all documents or other material that supported all such date(s).

THIRD SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 14

MIT objects to this Interrogatory as premature to the extent that it calls for a legal conclusion with respect to reduction to practice. MIT further objects to this Interrogatory to the extent that it mischaracterizes the legal standard for reduction to practice. MIT further objects to this Interrogatory because it seeks information protected by the attorney-client privilege, work product doctrine, and/or other applicable privileges or immunities.

Subject to and without waiving the foregoing general and specific objections, MIT states that the subject matter of the '685 patent was reduced to practice before the filing date of the application on which the '685 patent issued. The details of the reduction to practice were fully described in answer to numerous questions to the inventors propounded during the deposition testimony of Dr. James R. Davis, Ph.D., Christopher M. Schmandt, and MIT under Rule 30(b)(6) of the Federal Rules of Civil Procedure, and in response to this Interrogatory, those answers are herein incorporated by reference.

MIT states that the following claims were reduced to practice at least as early as June of 1989: 1-4, 7-9, 21, 23-25, 32-37, 40, 41, 53, 55, 57, and 58. MIT further states that the following claims were reduced to practice at least as early as August 4, 1989: 10-12, 15, 19, 20, 26-31, 42-49, 51, 52, and 54. MIT further states that the following claims were reduced to practice at least

as early as the filing date of the '685 patent, August 9, 1990: 5, 6, 13, 14, 16-18, 20, 22, 38, 39, 50, and 56.

MIT identifies documents bearing Bates numbers MIT00433-MIT00947, MIT01101-MIT01102, MIT01370-MIT01378, MIT01955-MIT02002, and MIT02155-MIT02274 as responsive to this Interrogatory. MIT further identifies pages 83-90, 104, 145-162, 179, and 287 of Mr. Schmandt's deposition transcript as responsive to this Interrogatory. MIT further identifies pages 79-91, 129, 167-169, 205, and 219-238 of Dr. Davis' deposition transcript as responsive to this Interrogatory. MIT further identifies pages 1-138 of MIT's 30(b)(6) deposition testimony as responsive to this Interrogatory.

Dated: June 16, 2006

Respectfully submitted,

Massachusetts Institute of Technology,

By its Attorneys,

Steven M. Bauer (BBO# 542531)
Kimberly A. Mottley (BBO# 651190)
Jacob K. Baron (BBO # 652568)
John W. Pint (BBO# 660548)
PROSKAUER ROSE LLP
One International Place
Boston, Massachusetts 02110-2600
Phone: 617-526-9600
Fax: 617-526-9899

CERTIFICATION

I, the undersigned, have reviewed MIT's Supplemental Responses to Harman's Interrogatory Nos. 1-5, 8, 11, and 13-14. The responses set forth herein, subject to inadvertent or undiscovered errors or omissions, are based on and therefore necessarily limited by the records and information still in existence, presently recollected, thus far discovered in the course of preparation of the responses, and currently available to MIT. Consequently, MIT reserves the right to make any changes in or additions to any of these responses if it appears at any time that errors or omissions have been made therein or that more accurate or complete information has become available. Subject to the limitations set forth herein, said responses are true to the best of my present knowledge, information and belief.

I hereby certify under penalty of perjury that the foregoing is true and correct.

Executed on this __th day of June, 2006.

John H. Turner, Jr.
Associate Director, Technology Licensing Office
On behalf of Massachusetts Institute of Technology

CERTIFICATE OF SERVICE

I HEREBY CERTIFY that on June 16, 2006, I caused a true and correct copy of MIT's SUPPLEMENTAL RESPONSES TO HARMAN'S INTERROGATORY NOS. 1-5, 8, 11 AND 13-14 to be served on the following counsel of record via email:

Robert J. Muldoon, Jr.
SHERIN AND LODGEN, LLP
101 Federal Street
Boston, MA 02110

William A. Streff Jr., P.C.
Michelle A. H. Francis
Craig D. Leavell
Ann H. Chen
Colleen M. Garlington
KIRKLAND & ELLIS LLP
200 East Randolph Drive
Chicago, IL 60601
(312) 861-2000 (phone)
(312) 861-2200 (fax)

John W. Pint

EXHIBIT
4

FILED UNDER SEAL

EXHIBIT

5

The Back Seat Driver: Real Time Spoken Driving Instructions

James Raymond Davis and Christopher M. Schmandt
The Media Laboratory
Massachusetts Institute of Technology
Voice: (617)-253-0314
Fax: (617)-258-6264

Abstract

The Back Seat Driver is an automobile navigation aid which uses synthetic speech to give driving instructions in real time to the driver of a car. The advantage of speech over visual aids is that it leaves the driver's eyes free for driving, however it also poses special problems. This paper describes the strategies employed by the Back Seat Driver to successfully use speech. We hope this paper will persuade you of the value of speech in driving directions.

Introduction

The Back Seat Driver uses synthetic speech to give driving instructions in real time to the driver of a car. Speech is the only output channel it uses. There are no graphics. This paper discusses the advantages and problems arising from our exclusive use of speech to provide directions. The first section presents a brief overview of the Back Seat Driver. The second section describes the linguistic abilities of the Back Seat Driver. The final section describes the problems we have encountered because of our exclusive use of speech, and how we have overcome them.

System Overview

The architecture of the Back Seat Driver is shown in figure 1. At the center of the Back Seat Driver is the map database. The street map represents two ways in which streets can be connected: *physical* connectivity means it is physically possible to drive from one segment to another, and *legal* connectivity means it is lawful to do so. Legal connectivity is obviously needed to find legal routes, and physical connectivity for correctly describing intersections. The street map also includes traffic lights, stop signs, the number of lanes, and the location of all gas stations. These features are useful for both route finding (since, e.g. fast routes should avoid traffic lights) and for descriptions. The location system (supplied by the project sponsor, NEC) determines the current position of the vehicle by dead reckoning and map matching. It is further described in [3]. The driver gives the Back Seat Driver a destination by entering an address on a keyboard.

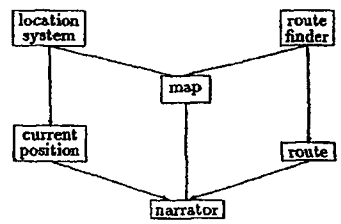


Figure 1: Back Seat Driver components

Using this map, the route finder can find the shortest route, the simplest one, or the one most easily followed, depending on the driver's preference.

The *narrator* is the subject of this paper. It generates instructions spoken by a speech synthesizer (a Dectalk). The narrator follows the driver's progress along the route. It decides what to say by comparing the current position against the map. The system follows the driver's progress, giving each instruction just when needed. If the time between instructions is long, the program gives the instruction twice, first in a detail, and later in a brief form. When not otherwise occupied, the system may deliver voice mail messages, weather reports, or commentary about the route. If the driver makes a mistake the system automatically finds an alternate route and continues.

The system has been running in prototype form since April 1989. It has been successfully used by drivers who have never driven in Boston. A somewhat longer description of the system appears in [4]. A complete description of the system appears in [1].

Linguistic Abilities

In designing the Back Seat Driver we chose to use speech as the sole means of providing driving instructions for two reasons. First, we believe that the driver's eyes are already employed watching traffic, and best left undisturbed. Second, we know that the alternative (map displays) will not work for those people who have difficulty reading maps[5]. We were also influenced by an experiment on route following which compared spoken instructions with paper maps[6]. Subjects who heard spoken directions did better than those with maps, and also better than those with *both* sources of guidance. Although this experiment does not compare real time speech to real time maps, it does suggest that spoken directions might be easier to follow than visual directions.

 CH2789-6/89/0000-0146$01.00 © 1989 IEEE

MIT 00938

Classifying Actions

Based on a study of how people naturally give spoken driving instructions, we developed a taxonomy of intersection types (Figure 2). This taxonomy is necessary in order to describe an intersection in the same way that a person would. For example, people talk about a "T" turn differently than a "fork" (or "Y") in the road. It is important that instructions match people's perceptions of the the world they see.

The proper classification of an intersection depends upon the topology (how many streets are at an intersection), the geometry (the angles among them), and the types of roads involved. For instance, the difference between the "T" and "fork" mentioned above is one of geometry, not topology (figure 3), and the difference between a "fork" and an exit from a highway is that one of the two roads in the "Y" of the exit is much larger than the other.

In our system, a route is a sequence of street segments leading from the origin to the destination. We consider every connection from one segment to another as an "intersection", even if there is only one next segment at the intersection. At any moment, the car will be on one of the segments of the route, approaching an intersection (unless an error occurs, which is handled as discussed below). The task of the Back Seat Driver is to say whatever is necessary to get the driver to go from the current segment, across the intersection, to the next segment of the route.

The items in the taxonomy of intersection types are called **acts**. We use an object oriented programming methodology, so for each act there is a corresponding "expert". The Back Seat Driver generates speech by consulting these experts. At any moment, there will be exactly one expert in charge of telling the driver what to do. To select this expert, the Back Seat Driver asks each expert in turn to decide whether it applies to the intersection. The experts are consulted in a fixed order, the most specific ones first. The first expert to claim responsibility is selected. This expert then has the responsibility of deciding what (if anything) to say.

- CONTINUE
- FORCED-TURN
- TURN-AROUND
- TURN
- FORK
- ENTER
- EXIT
- ONTO-ROTARY
- EXIT-ROTARY
- STOP

Figure 2: Act taxonomy

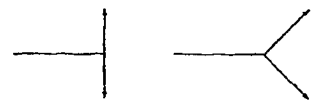


Figure 3: A "T" and a "Y" have the same topology

Describing actions

Each expert is able to generate text which describes the intersection. A description for an act must tell the driver two things: what to do and when (or where) to do it. "What to do" is expressed by a more or less constant verb phrase which depends upon the taxonomic classification, but may also depend upon specifics of the intersection. Thus a slight turn might be described by the verb "bear" where a sharper turn would be a "turn". The descriptions can be verbose or brief, and they can be expressed in past, present, or future tense. (We'll say why this flexibility is needed below.)

Saying "when"

Our study of natural instructions showed us that people almost never use distance as a cue for when to act. This is in sharp contrast to the textual directions provided by systems such as that of the Hertz rental company. Instead, people use two strategies. They wait until the driver is close to the intersection before saying anything, and/or they use a great variety of landmarks – including traffic lights, stop signs, other signs, buildings, road features, and the positions of other moving objects (e.g. "Follow that car."). The Back Seat Driver adopts both of these strategies.

Speech is especially useful as a cue for timing because speech is a temporal event, with a clear beginning and ending time. You know when someone begins to speak and when they finish. Someone peering at a map displayed on a CRT may have trouble distinguishing two adjacent streets, but there is no mistaking the word "now". Using time as a cue minimizes the workload on the driver, because the navigator absorbs the burden of remembering when to act. It also demands that the navigator have an accurate idea of where the car is. Our system demands positional accuracy of no greater than 10 meters for successful operation.

The Back Seat Driver's use of landmarks is unique in vehicle navigation systems. Our database began as a DIME file, but we extended it to include traffic lights, stop signs, road features (such as overpasses, bridges, and tunnels), distinctive signs, and the location of gas stations. Most of these are represented as attributes of the segments in the map database. To select a landmark for an intersection, the Back Seat Driver looks backwards from the intersection for the closest landmark which is also unique – that is, it prefers to say "take the first right after the underpass" rather than "take a right at the second set of lights". We think this makes the landmark easier to remember.

The Back Seat Driver does not speak at every single intersection. In the great majority of cases, it is perfectly obvious to the driver what to do (namely, to continue on forward). The action experts are also capable of deciding when the action at the intersection should be obvious to the driver. At present, the only action that is *ever* treated as obvious is CONTINUE. It is usually obvious to continue across an intersection, but we have found that what is obvious to one driver may not be so to another. Some people, for instance, are not comfortable driving across a major intersection unless they are instructed to do so. The expert can be somewhat customized so that its judgment of "obviousness" will correspond to that of the driver. If the action at the next intersection is obvious, the Back Seat Driver says nothing about it, and looks ahead for action at the next intersection, until it finds one that is *not* obvious.

The Back Seat Driver gives instructions just prior to the action. It also gives instructions further in advance, if time permits. This is especially useful when the instructions are complicated, as they are at some intersections. It is also able to give instructions "on demand". We call this the "what now" button. Drivers use this button for two reasons. Sometimes they are unsure whether they have come to the place where they are supposed to act, so they press the button to find out. At other times, they reach an intersection where the Back Seat Driver says nothing, because it believes the action is obvious, but it is not obvious to the driver. When the driver hits the "what now" button, the expert for the upcoming intersection describes it, even if it is considered to be obvious.

Talking about past and future

An advantage of language over pictures or gestures is that it can express events in the past or future. This advantage is well appreciated by readers of fiction, but may not yet be appreciated by designers of navigation systems. A navigation system should be able to talk about the past and future of the route, not just the present.

Drivers often need advance notice to prepare for an action. An example is what we call **lane advice**, which tells the driver to get into, or stay out of, a given lane. Lane advice is common in natural directions, and is one of the most appreciated features of the Back Seat Driver.

One reason for talking about the past is to describe mistakes. Drivers do not always follow the route the Back Seat Driver intends, either because of a mistake by the driver, the program, or external circumstances. When a mistake occurs, the Back Seat Driver finds a new route from the current location to the destination, while the driver is still moving. It also describes the mistake, saying something like "Oops, I meant for you to go straight." We think it is important that the system tell the user that there has been a mistake (without casting any blame on the user!) so that the user will come to better understand the system's style of instruction giving, and so that the user will remain confident in the system's understanding of the route. Talking about past and future actions is important in navigation. Speech seems to be the easiest way of doing this.

Example

As an example, here's a sample of the description of the left turn from Fulkerson Street to Main Street in Kendall Square, Cambridge.

> Get in the left lane because you're going to take a left at the next set of lights. It's a complicated intersection because there are two streets on the left. You want the sharper of the two. It's also the better of them. After the turn, get into the right lane.

This description was generated by the TURN expert in verbose form. It begins with some lane advice, then specifies the next action and provides a landmark for the place. The turn is described, and the proper street is described by two independent cues, one geometric, and one qualitative. Finally, the text provides a second piece of advice for after the turn.

Summary

The speech interface of the Back Seat Driver provides instructions without requiring the driver to look away from the road. Using speech permits us to talk about the past and the future as well as the present, and to give more detailed descriptions of the act than are possible with maps. Furthermore, it allows us to specify timing with great precision. But speech is not without its problems. The next section will discuss them, and the steps we have made to overcome them.

Liabilities of Speech

The advantages of a spoken language interface, as described above, do not come without cost. First, there are problems common to any natural language interface: while it is not terribly difficult to make a rudimentary interface, language generation requires substantial programming effort to be fluent and natural. Language is complicated, and people have literally a lifetime of experience with it, and are sensitive to fine nuances. On the other hand, having made this effort, we can exploit these nuances to to convey extra information.

A second problem is that a natural language interface is only useful to those who speak the language. In our experience, only a few non-native speakers have been able to understand the directions. Map displays have conventions of their own, but are more universal than natural language. We have also noticed that some driving terms used in the Boston area (e.g. "rotary") are not in the dialect of other English speakers. In our view, universality is not a prime concern. We believe that systems should be custom fit to the idiosyncrasies of their owners. The Back Seat Driver in your car should speak to you in the language and terms that are best for you as an individual, not you as a generic human.

The remainder of this section discusses problems specific to *spoken* natural language generation.

Speech takes time

As we said above, speech is inherently temporal. We take advantage of this when we use speech as a timing cue, but it also can be a difficulty. A real time spoken navigation system must plan its speech to ensure that it has enough time to say what it needs to say. If little time remains, it must say less (or speak more quickly), or ask the driver to slow down. We handle this problem by tracking the vehicle's position and velocity, and by modeling the time required to speak. The Back Seat Driver begins its speech at a time chosen to be early enough to allow the driver to hear the entire message, understand it, and react to it, before the point where action must be taken. The model of reaction time includes a constant for the driver's comprehension and a variable time which depends on the speed of the car, according to the maximum comfortable braking deceleration.

The temporal nature of speech also requires that the Back Seat Driver sometime combine instructions into a single utterance. When uttering an instruction, the Back Seat Driver looks ahead for the next instruction. If it determines that the time between the end of the execution of the current instruction and the beginning of the next is too short to allow it to speak the next instruction, it combines that text into the current one.

The Back Seat Driver does more than just give directions. Among other things, it also reads electronic mail messages from our office, gives weather reports, and makes comments about the route and road. Because speech takes time, and because a spoken utterance is only useful if completely spoken, the Back Seat Driver must carefully allocate the right to speak among potential tasks. It is undesirable for one task's speech to interrupt another's.

Speech can be misunderstood

A liability of speech, and synthetic speech in particular, is that speech can be misunderstood. This is particularly a problem with street names, because there are constraints that can help a driver correct a partially misunderstood name. A driver hearing an utterance that sounds like "Tarn reft" can guess that it is a corrupt form of "Turn left", but nothing can help the driver know what was intended by "Tarn Street". Directions should not use street names, because street name signs may be hard to see, misaligned, or simply missing. The importance of this first became apparent when we observed one driver who consistently misunderstood names, but also did not realize that he had misunderstood. Furthermore, the strength of his faith in the name was so strong that he drove straight through intersections, despite being told to "take the next left". This is probably the right thing to do with human instructions, where names are usually correctly understood, but street counts (e.g. "the third right") are imprecise or simply wrong. Our directions are phrased to minimize the use of street names in instructions. A typical text is: "Take the second left. It's Franklin Street."

Speech is transient

Information presented by speech does not persist, except in short term memory. We have already mentioned this as a reason why instructions should be given as late as possible. Another consequence of the transience of speech is that the system must be able to repeat itself at anytime, since the driver may not always be able to hear the speech. Repetition in turn poses a challenge.

since, unlike a program which reads the newspaper aloud, a literal repetition may not be appropriate, since the situation changes over time. For instance, if asked to repeat "Take the third left", the system may instead say "Take the second left" if the car has crossed an intersection. The consequence for the implementation is that the system retains not its previous words, but rather the previous reason for speaking. When asked to repeat, it invokes the same function that produced the last utterance.

A second problem with the ephemeral quality of speech is that the driver has no evidence of the program's existence except when it is speaking. We consider it very important that the driver have continued confidence that the program is running correctly, is aware of the driver's position and progress, and is "seeing" the world in the same way the driver does. We have devoted substantial effort to maintaining the illusion of co-presence.

In the introduction to this section, we said that the nuances of language could be used to convey much information. Co-presence is an idea communicated more by nuance than by explicit statement. (People would laugh if the system said "I'm right here with you." It sounds like something a therapist would say.) One way we indicate co-presence through nuance is by using **deictic** pronouns. Deictics are words that "point" at something. In English, we have four deictic pronouns: "this", "that", "these", and "those". The first two are singular, the second plural. The difference between "this" and "that" (and "these" and "those") is that "this" refers to something close. We use this in referring to landmarks. When the landmark is close, we use the proximal form (e.g. "these lights"); when distant, we use a brief noun phrase (e.g. "the next set of lights"). This is important. When a driver is stopped 30 meters back from a stop light, it may be literally true to say "turn left at the next set of lights", but it will confuse the driver.

A second means of conveying co-presence is to acknowledge the driver's actions. After the driver carries out an instruction the system briefly acknowledges the act if there is time, and if the act was not so simple (e.g. continuing straight) as to need no acknowledgment. This acknowledgment is a short phrase like "Okay". Some drivers dislike acknowledgments, so they can be disabled, but most find the confirmation comforting. The timing of the acknowledgment does much to confirm the driver's sense that the program really knows where the car is. Another source of acknowledgment is the use of **cue words** in the instructions. It will often by the case that the route calls for the driver to do the same thing twice (e.g. make two left turns). The speech synthesizer we use has very consistent pronunciation, and drivers sometimes get the impression that the system is

repeating itself because it is in error (like a record skipping). The acknowledgments help to dispel this, but we also cause the text to include cue words such as "another". These indicate that the system is aware of its earlier speech and the driver's previous actions.

Yet another means of conveying co-presence is to make occasional remarks about the road and the route. These remarks indicate that the program is correctly oriented. As an example, when the road makes a sweeping bend to one side, the program speaks of this as if it were an instruction ("Follow the road as it bends to the right.") even though the driver has no choice in what to do. The program also warns the driver about potentially hazardous situations, such the road changing from one-way to two-way, or a decrease in the number of lanes. As with acknowledgments, these warnings can be disabled if the driver dislikes them. Other remarks have less to do with the route. We justify these by the maxims of cooperative conversations formulated by philosopher H. P. Grice[2]. His maxim of QUANTITY (part 1) says: "Make your contribution as informative as is required." Grice explains that one can convey information by appearing to flout the maxim. In this case, a driver can reason as follow: "The program, like all cooperative agencies, obeys the maxim of quantity. Therefore, it is had something important to say, it would say it. The program said nothing of great significance, therefore there is nothing urgently requiring my attention. So everything is well." At present, our "Gricean" utterances are trivial observations about the weather, but we are re-designing them to convey useful information about the city.

Summary

A speech interface for giving driving instructions has several advantages over a graphics interface. There are problems with natural language interfaces in general, and speech in particular, but they can all be overcome. The result is an excellent aid for navigation.

Acknowledgments

The authors wish to gratefully acknowledge the support of NEC Home Electronics, Ltd.

References

[1] James Raymond Davis. *Back Seat Driver: voice assisted automobile navigation.* PhD thesis, Massachusetts Institute of Technology, September 1989.

[2] H. P. Grice. Logic and conversation. In Cole and Morgan, editors, *Syntax and Semantics: Speech Acts*, volume 3, pages 41–58. Academic Press, 1975.

[3] Osamu Ono, Hidemi Ooe, and Masahiro Sakamoto. CD-ROM Assisted Navigation System. In *Digest of Technical Papers*, pages 118–119. IEEE ICCE, 1988.

[4] Christopher M. Schmandt and James R. Davis. Synthetic speech for real time direction giving. *IEEE Transactions on Consumer Electronics*, page (to appear), 1989.

[5] Lynn A. Streeter and Diane Vitello. A profile of drivers' map reading abilities. *Human Factors*, 28:223–239, 1986.

[6] Lynn A. Streeter, Diane Vitello, and Susan A. Wonsiewicz. How to tell people where to go: comparing navigational aids. *International Journal of Man/Machine Systems*, 22(5):549–562, May 1985.

EXHIBIT

6

FAM 19.2

Synthetic Speech For Real Time Direction-Giving

Christopher M. Schmandt and James Raymond Davis
The Media Laboratory
Massachusetts Institute of Technology

Abstract

The Back Seat Driver is a research prototype of a system to use speech synthesis as a navigational aid for an automobile equipped with localization equipment. We are evaluating the user interface by field trials. As this is work in progress, this paper will primarily give an overview of the system and describe its components. Included will be discussion of the map database, route finding algorithm, repair strategies, and the discourse generator.

Goals

The main goal of this project is to evaluate the utility of speech synthesis as the user interface to a real-time navigation system in an urban environment. Of particular concern is the discourse structure:

- how should driving acts be described?
- how can a description be generated from a route?
- how should timing considerations be applied?
- what kinds of feedback, both positive and negative, does the user require?
- what kinds of visual cues are most useful in describing an approaching location?

This information is gained from both laboratory simulations and field trials.

Our desire is to build the best possible real-time route describer. Although we believe a speech interface to the navigation unit is superior and safer than a visual interface, we do not plan to conduct direct comparison studies.

In the course of field trials to evaluate and improve our automatic direction giving, we hope to specify key components of the map database. We expect discourse behavior may need to vary with conditions (traffic, weather, day/night). It is likely that different visual cues may be useful in these situations. All must be represented in the database.

Geographic Database

Our database covers 41 square miles in the Boston area, including parts of Boston, Cambridge, Brookline, Somerville, and Watertown. It originated as a DIME (Dual Independent Map Encoding) file distributed by the United States Geological Survey[1]. A DIME file consists of a set of straight line segments, each with a name, a type, endpoints in longitude and latitude, and some additional information such as address numbers. Initially our database contained many errors. Correcting them required actually traveling most of the segments.

A DIME file alone is not sufficient for finding routes. The DIME files indicate physical connectivity, but route finding requires *legal* connectivity, i.e., one can legally drive from one segment to the next (one way streets are a simple example). We extended the data base format to explicitly represent legal connectivity. Since some streets are better than others, the database must include a measures of *quality*. We take this to be a largely subjective measure of the ease of locating and following a street. This allows the route finder to prefer to use streets of higher quality.

The generation of easy followed natural descriptions requires more extensions. We added a number of new segment types to distinguish bridges, underpasses, tunnels, rotaries, and access ramps. All these extensions were done for an earlier route finding project[2].

We are presently adding landmarks to the database. Drivers need landmarks to know how far to drive and when to turn. If the Back Seat Driver had eyes, it could simply choose landmarks as needed by looking for them in the landscape. Being blind, it must rely on landmarks coded into the map database. We have already added traffic lights to the landmark database. A main task now is to determine what else must be added.

System

Our vehicle is equipped with a localization unit built by NEC Home Electronics, Ltd., the project sponsor. It is a dead-reckoning position keeping system which uses speed and direction sensors. To compensate for error, it uses map matching on a map database stored on CD ROM. The system described more fully in [3].

As this is a research prototype, much of the computation is done in a base station computer laboratory (on a Symbolics Lisp Machine), rather than a computer on the vehicle. Two cellular

MIT 01101

CH2724-3/89/0000-0288 $1.00 © 1989 IEEE

telephones link the computer to the car. The on board navigational hardware transmits position and velocity via modem and cellular phone to the base station. The base station computer does all route planning and discourse generation. Speech synthesis is performed in a commercial text-to-speech synthesizer (Dectalk) cabled to the Lisp Machine. Synthesized instructions to the driver are relayed via the second cellular link and a speaker phone in the car. The keypad of the second phone also serves as the driver's control unit for the Back Seat Driver. Through this keypad a driver selects a destination, requests repeats of spoken information, and accesses other services of the Back Seat Driver.

A block diagram of the system appears in Figure 2, below.

Discourse Strategies

The instructions are detailed and natural, and include a rich taxonomy of driving verbs. The dialog system uses cues such as vehicle speed and difficulty of driving actions to attempt to deliver instructions at the proper pace and in a timely manner. In addition, the system can anticipate some of the driver's possible mistakes and give warnings to avoid them.

If the driver does make a wrong turn, or misses a turn, the Back Seat Driver describes the error and then incrementally calculates a new route, rather than simply back-tracking to the point of the error. Route planning includes weighting for length of the trip, difficulty of driving maneuvers (such as left turns against traffic), street quality, and complexity of the spoken directions.

As opposed to much prior work in discourse generation, the Back Seat Driver is a real-time system which must factor in a number of temporal considerations. It needs to give each stage in the directions at just the right point, in terms of the time it takes to execute the driving maneuver as well as the speed of the vehicle approaching the intersection. For safety considerations, we would rather err on the side of giving the driver plenty of warning, but a cue given too far in advance may be miscued (e.g., a turn taken at an earlier intersection). Additionally, the software must consider the length of time it will take to recite an utterance. It is better to miss a turn and plan a new route than start describing the turn at a time when it may be unsafe to execute it (i.e., already well into an intersection).

Summary

The Back Seat Driver is already working in prototype form. Our present concerns are to determine what a spoken driving assistant should say, to understand how time and speed affect this decision, and to learn what features a map database must have to support generation of instructions.

Acknowledgments

The authors wish to gratefully acknowledge the support of NEC Home Electronics, Ltd.

References

[1] *Geographic Base File GBDF/DIME: 1980 Technical Documentation.* U.S. Department of Commerce, Data Users Services Division, 1980.

[2] James R. Davis and Thomas F. Trobaugh. *Direction Assistance.* Technical Report 1, MIT Media Laboratory Speech Group, Dec 1987.

[3] Osamu Ono, Hidemi Ooe, and Masahiro Sakamoto. CD-ROM Assisted Navigation System. In *Digest of Technical Papers,* pages 118–119, IEEE ICCE, 1988.

Figure 1: Map of Boston area

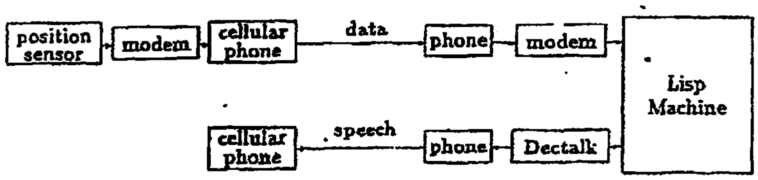


Figure 2: Block diagram of communications

EXHIBIT

7

Excerpts from the:

May 19, 2006
30(b)(6) Deposition of
Christopher Schmandt

Volume: I

Pages : 1 - 140

Exhibits: 90 - 96

UNITED STATES DISTRICT COURT

DISTRICT OF MASSACHUSETTS

CIVIL ACTION NO. 05-10990-DPW

\- -

MASSACHUSETTS INSTITUTE OF TECHNOLOGY,

Plaintiff,

V.

HARMAN INTERNATIONAL INDUSTRIES INCORPORATED,

Defendant.

\- -

CONFIDENTIAL

VIDEOTAPED 30(b)(6)DEPOSITION OF M.I.T.

through CHRISTOPHER SCHMANDT

Friday, May 19, 2006, 2006, 9:40 a.m.

Proskauer Rose LLP

One International Place

Boston, Massachusetts

Reporter: Rosemary F. Grogan, CSR, RPR

LegaLink Boston, a Merrill Company

APPEARANCES:

Representing the Plaintiff:

PROSKAUER ROSE LLP
One International Place
Boston, MA 02110
(617)526-9700
kmottley@proskauer.com
BY: KIMBERLY A. MOTTLEY, ESQUIRE

Representing the Defendant:

KIRKLAND & ELLIS LLP
200 East Randolph Drive
Chicago, IL 60601
(312)861-2105
cleavell@kirkland.com
BY: CRAIG D. LEAVELL, ESQUIRE

Also present:

Jason Moschella, Videographer

09:34:12 THE VIDEOGRAPHER: This is the beginning of videocassette No. 1 in the deposition of the Massachusetts Institute of Technology by and through Christopher Schmandt in the matter of M.I.T., plaintiff, versus Harman International Industries Incorporated, defendant, in the United States District Court, District of Massachusetts, Civil Action No. 05-10990-DPW.

09:40:11 Today is May 19, 2006. The time is 9:40 a.m. My name is Jason Moschella. I'm a certified legal video specialist and a notary public contracted by LegaLink Boston. This deposition is taking place today at the offices of Proskauer Rose LLP, One International Place, Boston, Massachusetts and was noticed by Kirkland & Ellis.

09:40:35 At this time counsel will please identify yourselves and the court reporter will administer the oath.

09:40:42 MR. LEAVELL: This is Craig Leavell from Kirkland & Ellis on behalf of Harman International.

09:40:44 MS. MOTTLEY: This is Kimberly Mottley of Proskauer Rose on behalf of M.I.T.

CHRISTOPHER SCHMANDT, having been satisfactorily identified by the production of a driver's license, and duly sworn by the Notary Public, was examined and testified as follows:

EXAMINATION BY MR. LEAVELL:

Q. Good morning, sir.
A. Good morning.
MR. LEAVELL: I'm going to ask the court reporter to mark the next exhibit as Exhibit 90.
(Exhibit 90 Marked for Identification)
BY MR. LEAVELL:
Q. Mr. Schmandt, have you seen a copy of Exhibit 90 prior to today, which is the Notice of Harman's First Rule 30(b)(6) Deposition of MIT, for the record?
A. Yes, I have.
Q. And you've had a chance to consult with counsel for M.I.T. about the scope of today's deposition; is that correct?
A. Yes, I have.
MS. MOTTLEY: I can represent for the record, Craig, that Mr. Schmandt is ready to testify on

09:46:54 A. If I'm not mistaken, this document is from June of 1989.
09:47:01 Q. You're right. Thank you for the correction.
09:47:05 Were there field trials of the Back Seat Driver that took place prior to June of 1989?
09:47:11 A. One would certainly infer that from the sentence in the paper, so I believe that to be true.
09:47:17 Q. And on behalf of M.I.T. it's your 30(b)(6) testimony that is true, correct?
09:47:24 A. That's correct.
09:47:24 Q. And what about in July of 1989, did field trials continue of the Back Seat Driver System in July of 1989?
09:47:32 A. It's highly likely they continued in July of 1989. There's no specific record.
09:47:37 Q. Is there any reason to believe that field trials of the Back Seat Driver did not continue into July of 1989?
09:47:45 A. No, there is no reason to believe they did not continue. Perhaps we should discuss this term, Field Trials, however. This may be the cause of some confusion here.
09:47:55 In some industries, for example, telecommunications, a cellular phone provider might do

09:48:03 what is called a field trial of a service, which is
09:48:06 basically make the service available to its customers in
09:48:09 some marketplace to test whether they want to do a
09:48:13 national roll-out, for example. In our case, field
09:48:17 trials, for the purposes of this work, is not being used
09:48:20 that way.
09:48:21 A field trial is any trial that occurs
09:48:24 outside the building.
09:48:25 Q. Okay. So there were instances in June of
09:48:29 1989, in which the Back Seat Driver System was used
09:48:34 outside of the Media Lab driven on a public street in
09:48:39 the Boston area, correct?
09:48:40 A. Could you say that again, the beginning of
09:48:42 that question?
09:48:43 Q. Sure.
09:48:43 In June of 1989, there were field trials
09:48:46 that included using the Back Seat Driver System on
09:48:51 public streets in the Boston area; is that correct?
09:48:55 A. There is no specific record of whether any
09:48:58 trials occurred during the month of June.
09:49:00 Q. Is there any reason to believe that the Back
09:49:02 Seat Driver was not driven on public Boston streets
09:49:07 during the time frame of June of 1989?
09:49:10 A. There's no reason not to believe that.

09:49:12 Q. Is there any reason to believe that the Back Seat Driver System was driven on public streets in the Boston area in July of 1989?
09:49:24 A. Again, my answer would be the same.
09:49:26 Q. There's no reason not to believe that; that is true?
09:49:33 A. That's correct.
09:49:41 Q. So is it reasonable to believe that there were indeed uses of the Back Seat Driver on public roads in Boston in June of 1989?
09:49:53 A. I think I've already answered that question.
09:49:55 Q. You've answered it yes, correct?
09:49:56 A. No, I've answered that there's no reason to believe that there wasn't. I have no evidence that there were trials; that there were driving in the month of June in 1989.
09:50:09 Q. Is it reasonable to believe that there were -- that somebody during the month of June in 1989, drove the Back Seat Driver on a public street in Boston in June of 1989?
09:50:22 A. I think that's the same question I answered before, and I think the answer was yes.
09:50:27 Q. What about July of 1989?
09:50:29 A. The answer again would be the same.

EXHIBIT

7

(Part 2)

Q. That it's reasonable to believe that somebody was driving on public Boston streets in July of 1989 using the Back Seat Driver System?

A. At least once, yes, it's likely.

Q. Is it also likely to believe -- or is it likely and reasonable to believe that somebody was driving the Back Seat Driver on a public road in the Boston area in the month of August of 1989?

A. I think it's highly unlikely.

Q. Why do you say that?

A. Because in the month of August, Jim, who would have been the principal researcher -- I'm sorry, the researcher would have been in the car in the course of doing those trials, those drivings, was working very hard to finish his thesis. And it is certainly possible that driving would have occurred, but I think it's highly unlikely because I think at that point he had a very large document to produce.

Q. Let's talk about the system as it existed in June and July of 1989.

First of all, is there any reason to believe that any changes to the Back Seat Driver took place between June and July of 1989 or can we talk about that as one system?

10:05:41 his thesis.
10:05:42 Nonetheless, in June, we certainly -- we
10:05:48 both would have predicted with relative confidence that
10:05:51 Jim's thesis was going to be finished at the end of the
10:05:54 summer.
10:06:13 Q. And the summary on Exhibit 91, it says, this
10:06:18 is on the second page, it says: The Back Seat Driver is
10:06:32 already working in prototype form.
10:06:37 Was that a true statement? Was the Back
10:06:39 Seat Driver working in prototype form as it existed in
10:06:43 June of 1989?
10:06:45 MS. MOTTLEY: Objection; outside the scope.
10:06:46 A. Yes, it was.
10:06:51 Q. And it was already working as a prototype
10:06:56 system that included each of the limitations in Claim 1
10:06:59 of the '685 patent in June of '89; is that correct?
10:07:02 MS. MOTTLEY: Objection; outside the scope.
10:07:03 A. Yes, that's correct.
10:07:16 Q. What about Claim 2 of the '685 patent, did the
10:07:20 Back Seat Driver, as it existed in June of '89 and was
10:07:24 working in prototype form in field trials, include the
10:07:34 subject matter recited in Claim 2 of the '685 patent?
10:07:38 MS. MOTTLEY: Objection as outside the scope
10:07:39 and also objection to the extent that you're asking

about claim terms that the Court has not yet construed, and thus the question is vague.

You can answer.

A. Yes, it did.

Q. How do you know that?

A. Because that was the form of the Back Seat Driver database.

Q. When did the Back Seat Driver database take that form?

MS. MOTTLEY: Same objections.

A. I can't tell you any particular date --

Q. Prior to --

A. -- at which that existed.

Q. Prior to June of '89?

A. In all --

MS. MOTTLEY: Objection.

A. In all of the test drives of the Back Seat Driver, the database took that form.

Q. Including the test drives that took place in July or early August of '89?

MS. MOTTLEY: Same objections.

A. The term that I used was all.

Q. So yes?

A. Yes.

10:08:27 Q. Thank you.
10:08:31 What about Claim 3 in the '685 patent,
10:08:35 did the Back Seat Driver, as it existed in the
10:08:40 successful field trials in June of '89, include the
10:08:44 subject matter recited in Claim 3 of the '685 patent?
10:08:47 MS. MOTTLEY: Same objections.
10:08:49 A. I'm concerned because suddenly you've
10:08:52 introduced this word, successful. You refer to
10:08:56 successful field trials.
10:08:56 Q. Let me rephrase --
10:08:57 A. There was not any particular metric of success
10:09:00 in these.
10:09:01 Q. Let me rephrase.
10:09:03 Was the Back Seat Driver, that was
10:09:03 already working in prototype form in field trials in
10:09:08 June of '89, include the subject matter recited in Claim
10:09:11 3 of the '685 patent?
10:09:14 MS. MOTTLEY: Same objections.
10:09:15 A. Yes, it did.
10:09:16 Q. How do you know that?
10:09:17 A. Because all --
10:09:18 MS. MOTTLEY: Same objections.
10:09:19 THE WITNESS: Sorry.
10:09:19 A. Because all field trials of the Back Seat

10:09:20 Driver met that claim.
10:09:24 Q. Including field trials that took place in July
10:09:27 or August of '89?
10:09:29 A. Yes.
10:09:30 MS. MOTTLEY: Same objections.
10:09:30 A. As before, that's what I mean by all.
10:09:33 Q. Thank you.
10:09:34 What about Claim 4 of the '685 patent,
10:09:38 did the Back Seat Driver that was working in prototype
10:09:41 form in field trials for June of '89 include the subject
10:09:44 matter recited in Claim 4 of the '685 patent?
10:09:49 MS. MOTTLEY: Same objections.
10:09:50 A. Yes, it did.
10:09:50 Q. How do you know that?
10:09:51 MS. MOTTLEY: Same objections.
10:09:52 A. Because as before, all field trials use a
10:09:56 database of that form.
10:09:57 Q. Including in July and early August of '89?
10:10:00 MS. MOTTLEY: Same objection.
10:10:02 A. As before all includes July and August.
10:10:09 Q. Did the Back Seat Driver that was already
10:10:12 working in prototype form in June of 1989 field trials
10:10:20 include the subject matter recited in the Claim 5 of the
10:10:23 '685 patent?

11:24:44 Q. And how do you know that?

11:24:48 A. Because that's how the software was written.

11:24:52 Q. From the very beginning of the Back Seat Driver?

11:24:54 MS. MOTTLEY: Same objections.

11:24:55 A. Yes.

11:25:01 Q. So all subsequent trials after June of '89, also included the subject matter of Claim 41, correct?

11:25:07 MS. MOTTLEY: Same objections.

11:25:08 A. Yes.

11:25:14 Q. And the Back Seat Driver that was an actual working prototype that had successfully guided drivers unfamiliar to Cambridge to their designations prior to August 4th of 1989, also included the subject matter of claim 41 of the '685 patent, correct?

11:25:29 MS. MOTTLEY: Same objections.

11:25:31 A. Yes.

11:25:44 Q. Now, your list of Exhibit 92 has all of claims 42 through 49 under the Yes category, correct?

11:25:52 A. Yes.

11:25:52 Q. Does that mean that the subject matter of each of Claims 42 through 49 was present in the Back Seat Driver that had been an actual prototype and had successfully guided drivers unfamiliar with Cambridge to

11:26:11 their designations on the public streets around Boston
11:26:15 prior to August 4th of 1989, did that system at that
11:26:19 time include the subject matter of each of Claims 42
11:26:23 through 49 inclusive?
11:26:25 MS. MOTTLEY: Same objections.
11:26:26 A. Yes.
11:26:34 Q. Did the Back Seat Driver, as it existed as a
11:26:39 working prototype in the field trials in June of '89,
11:26:40 include the subject matter of Claims 42 through 49
11:26:44 inclusive?
11:26:46 MS. MOTTLEY: Same objections --
11:26:46 A. I haven't --
11:26:47 MS. MOTTLEY: -- and compound.
11:26:48 A. I haven't considered that before. I'll have
11:26:51 to study the claims if you would like me to answer it.
11:26:54 Q. Please take a moment, or however much time you
11:26:56 need, to answer that question.?
11:27:44 MS. MOTTLEY: You're asking as to 42 through
11:27:47 49 inclusive?
11:27:48 MR. LEAVELL: Yes. That's what the current
11:27:50 pending question is. I'm happy to break it down
11:27:54 and do it separately --
11:27:56 MS. MOTTLEY: It might be easier.
11:27:57 MR. LEAVELL: -- or we can separate it and do

11:28:00 it through 42 through 47 and deal with 48 and 49,
11:28:04 but whatever the witness needs to do.
11:28:06 MS. MOTTLEY: Sure.
11:28:35 A. Okay. To make sure I'm answering the proper
11:28:37 question, the question was as of June of 1989?
11:28:42 Q. Right. The question is: Did the Back Seat
11:28:44 Driver, as it existed as a working prototype in field
11:28:48 trials in June of 1989, include the subject matter of
11:28:53 Claims 42 through 49?
11:28:56 A. We don't know.
11:29:02 Q. Why not?
11:29:02 A. Because we have no documentation that tells us
11:29:06 when those features were added to the system.
11:29:25 Q. If those features were included in the
11:29:29 database of the Direction Assistance Program, is it
11:29:35 reasonable to assume they were carried over to the Back
11:29:37 Seat Driver System?
11:29:38 MS. MOTTLEY: Same objections.
11:29:39 A. No, it is not.
11:29:41 Q. Why not?
11:29:42 MS. MOTTLEY: Same objections.
11:29:43 A. Because these aren't -- these claims don't
11:29:48 apply to databases. They apply to what's spoken.
11:29:52 Q. But the Back Seat -- or the Direction

11:29:54 Assistance provided spoken instructions that classified the instructions into intersection types, correct?

11:30:02 MS. MOTTLEY: Same objections.

11:30:04 A. I don't have an answer on that. I'm not prepared today to talk about the Direction Assistance.

11:30:19 Q. Did M.I.T. ever make a navigation system that included the subject matter of Claim 50 of the '685 patent?

11:30:28 MS. MOTTLEY: Same objections.

11:30:34 A. No, it did not.

11:31:05 Q. Did the Back Seat Driver that was an actual working prototype and that had successfully guided drivers unfamiliar with Cambridge to their designations prior to August 4th of 1989, did that system include the subject matter of Claim 51 of the '685 patent?

11:31:26 MS. MOTTLEY: Same objections.

11:31:35 A. Yes.

11:31:36 Q. And you know that because the thesis says so?

11:31:38 A. Yes.

11:31:40 Q. Did the Direction Assistance, as it existed as a working prototype in field trials in June of '89, include the subject matter of Claim 51?

11:31:52 MS. MOTTLEY: I objection. I think you said Direction Assistance.

14:41:42 SIGNATURE / ERRATA SHEET
14:41:42 Re: M.I.T Vs. Harman International Industries
14:41:42 DEPOSITION OF: Christopher Schmandt 5/19/06
14:41:42 I, CHRISTOPHER SCHMANDT, do hereby certify
14:41:42 that I have read the foregoing transcript of my
14:41:42 testimony, and I further certify that said transcript it
14:41:42 is a true and accurate record of said testimony (with
14:41:42 the exception of the corrections that are noted below).
14:41:42 PAGE LINE(S) READS SHOULD READ
14:41:42 ____
14:41:42 ____
14:41:42 ____
14:41:42 ____
14:41:42 ____
14:41:42 ____
14:41:42 ____
14:41:42 Signed under the pains and penalties of
14:41:42 perjury this ________day of ______________, 2006.
14:41:42 ____
14:41:42 CHRISTOPHER SCHMANDT Date
14:41:42 Subscribed and sworn to before me this ____day
14:41:42 of __________, 2006.
14:41:42 ____________________
14:41:42 Notary Public My Commission Expires:__________

EXHIBIT

8

FILED UNDER SEAL

EXHIBIT

9

Excerpts from the:

February 16, 2006
Deposition of
James R. Davis

VOLUME: I

PAGES: 1-248

EXHIBITS: 66-89

UNITED STATES DISTRICT COURT

DISTRICT OF MASSACHUSETTS

- - - - - - - - - - - - - - - - - - x

MASSACHUSETTS INSTITUTE OF TECHNOLOGY,

Plaintiff,

v. Civil Action No. 05-10990-DPW

HARMAN INTERNATIONAL INDUSTRIES INCORPORATED,

Defendant.

- - - - - - - - - - - - - - - - - - x

CONFIDENTIAL

VIDEOTAPED DEPOSITION of JAMES R. DAVIS

February 16, 2006

9:17 a.m.

Proskauer Rose

One International Place

Boston, Massachusetts

Reporter: Michael D. O'Connor, RPR

APPEARANCES:

PROSKAUER ROSE LLP
By Steven M. Bauer, Esq.
One International Place
Boston, Massachusetts 02110
(617)526-9700
For the Plaintiff.

KIRKLAND & ELLIS LLP
By Craig D. Leavell, Esq.
200 East Randolph Drive
Chicago, Illinois 60601
(312)861-2105
For the Defendant.

Also Present: Richard Mendes, Videographer

P R O C E E D I N G S

VIDEOGRAPHER: Here begins videotape number one in the deposition of James R. Davis in the matter of Massachusetts Institute of Technology versus Harman International Industries, in U.S. District Court of Massachusetts, Case No. 05-1099-DPW.

Today's date is February 16, 2006. The time on the video monitor is 9:17 a.m. The video operator today is Richard Mendes, contracted by LegaLink, Boston, Massachusetts. This video deposition is taking place at Proskauer Rose, One International Place, Boston, Massachusetts.

Counsel, please voice identify yourselves and state whom you represent.

MR. BAUER: Steven Bauer from Proskauer Rose representing M.I.T. and the witness.

MR. LEAVELL: Craig Leavell from Kirkland & Ellis representing the Defendant Harman.

VIDEOGRAPHER: The court reporter today is Michael O'Connor of LegaLink Boston. Would the reporter please swear in the witness.

JAMES R. DAVIS

having been satisfactorily identified by the production of his driver's license, and duly sworn by the Notary Public, was examined and testified as follows:

VIDEOGRAPHER: Please begin.

DIRECT EXAMINATION

BY MR. LEAVELL:

Q. Good morning, sir.

A. Good morning.

Q. I'm going to hand you a copy of what's been previously marked as Defendant's Exhibit 32, a copy of U.S. Patent No. 5,177,685. If I refer to that as the '685 patent or your patent, will you understand that that's what I'm referring to today?

A. I will.

Q. You are the James R. Davis that's listed as an inventor on the face of the '685 patent?

A. I am.

Q. Mr. Davis, who do you currently work for?

A. The Ontario Principals Council of Toronto,

demonstrated "The Back Seat Driver" of system at any of these conferences, like the one in June of '89?

A. I do not think it would have been possible for him to demonstrate the system, since to do that he would have to bring the car with him, and that wouldn't be easy to do.

Q. Well, a videotape was made of "The Back Seat Driver" system, correct?

A. I believe the video -- there was a videotape made. I believe it was made subsequent to September of '89.

Q. Why do you believe that?

A. I'm speculating now. I think it would be better if I saw the videotape.

Q. On the bottom of the front page of Exhibit 80, there's a date that says, "Manuscript received June 9, 1989." Do you see that?

A. I do.

Q. Do you have any reason to doubt that a manuscript for this article was supplied to the IEEE back in June of 1989?

A. I have no reason to doubt that.

Q. Is there any reason to believe that the IEEE was told that they had to keep that manuscript

in confidence after having received it in June of
'89?

A. "Confidence," that's a subtle word in the context of an academic conference. On the one hand, there is no obligation of confidentiality of the sort that, say, an attorney and her client have.

On the other hand, it would be a violation of the norms of academia were they to show the paper to anyone other than people whose immediate job was to assess the paper for quality of publication.

These are reviewed conferences, and it's the duty of the conference organizers to show or at least to attempt to assure themselves that the work is novel, isn't plagiarized, is relevant to the conference, that the paper is actually written in a way that the people can understand it, and there are people whose job is to then review the paper. I myself have served as such a reviewer.

Reviewers are -- I have no idea whether the reviewers in this case signed any contract or document stating explicitly the understanding, it is well known in academia that when you get such a paper, you are not to show it to anyone, other than those immediately concerned with reviewing it, until

Q. Did "The Back Seat Driver" as it existed in June of '89 as it was being used for field trials include in the database, traffic lights, stop signs and some buildings?

A. Yes.

Q. Did "The Back Seat Driver" as it existed in June of '89 and which was being used for field trials include lane information?

A. By "lane information," you mean?

Q. Both the number of lanes, as well as any turn restrictions on lanes, e.g., left turn only?

A. Yes.

Q. Did "The Back Seat Driver" as it existed in June of '89 that was being used for field trials include wording of signs in the database so that "The Back Seat Driver" can give directions by referring to the signs?

A. Yes.

Q. Did "The Back Seat Driver" as it existed in June of 1989 and as it was being used for field trials at that time include a computing apparatus for running and coordinating system processes?

A. Yes.

Q. Did it include driver input means

functionally connected to the computing apparatus for entering data into the computing apparatus, including a desired destination?

A. Yes.

Q. Did "The Back Seat Driver" system as it existed in June of '89 and as it was being used for field trials include a map database functionally connected to the computing apparatus which distinguished between physical and legal connectivity?

A. Yes.

Q. Did that same system as it existed then as it was being used for field trials include position sensing apparatus installed in the automobile and functionally connected to the computing apparatus for providing said computing apparatus data for determining the automobile's current position?

A. Yes.

Q. Did "The Back Seat Driver" as it existed in 1989 as it was being used for field trials include each of the limitations of claim one of the '685 patent?

A. I'm not sure I understand that question well enough to answer it. Would you be more

specific, please?

Q. Sure. Did "The Back Seat Driver" as it existed in June of 1989 as it was being used for field trials include a location system?

A. Do you mean a location system functionally connected to said computing apparatus?

Q. Sure. Did it include that?

A. Yes.

Q. Was that location system in that system as it existed at that time used for accepting data from the position-sensing apparatus and for consulting the map database and for determining the automobile's current position relative to the map database?

A. Yes.

Q. Did "The Back Seat Driver" as it existed in June of 1989 as it was being used for field trials include a route finder functionally connected to the computing apparatus?

A. Yes.

Q. Did that route finder accept the desired destination from the driver input means and the current position from the location system for consulting the map database and for computing a

route to the destination?
A. Yes, it did.
Q. Did "The Back Seat Driver" as it existed in June of '89 and was being used for field trials include a discourse generator functionally connected to the computing apparatus for accepting the current position from said location system and from the route finder for consulting the map database and for composing discourse, including instructions and other messages for directing the driver to the destination from the current position?
A. Yes.
Q. Did "The Back Seat Driver" as it existed in June of '89 as it was being used in field trials include a speech generator functionally connected to the discourse generator for generating speech from said discourse provided by said discourse generator?
A. Yes.
Q. Did that same system at that same time use a voice apparatus functionally connected to the speech generator for communicating the speech provided by the speech generator to the driver?
A. Yes.
Q. Did "The Back Seat Driver" as it existed in

June of '89 as it was being used for field trials
include a map database that had three-dimensional
representation of street topology?
A. Yes.
Q. So the map database in "The Back Seat
Driver" was a 3-D system at that time in June of
'89?
MR. BAUER: Objection. Vague.
A. My recollection on the nature of the 3-D
database isn't very clear at this time.
Q. Did "The Back Seat Driver" system as it
existed in June ever '89 and as it was being used
for field trials at that time, did the map database
in that system distinguish divided streets?
A. Yes.
Q. Did the map database in that system include
speed limits?
A. I don't recall.
Q. Did the map database in "The Back Seat
Driver" as it existed in June of '89 during the time
of the field trials include within the database the
expected rate of travel?
A. I don't recall.
Q. Did the map database in "The Back Seat

Driver" as of June of '89, while it was being used for field trials, include locations for service stations?

A. I don't recall.

Q. Did "The Back Seat Driver," as it existed in June of '89 and as it was being used for field trials at that time, include a best first route-searching algorithm?

A. I'm sorry, would you please repeat the question?

Q. Sure. What type of route-finding algorithm was used in "The Back Seat Driver" as it existed in June of '89 during the time of the field trials?

A. At that time we were using either the A* or a modified A* algorithm.

Q. Is the A* algorithm a type of best first search algorithm?

A. It is an improvement on best first. If you want to go into the mathematics of it, we can have a more lengthy explanation, if you'd like.

Q. In your mind, is it accurate to say that the A* algorithm is a type of best first search algorithm?

A. No.

determined the heading orientation of the vehicle?
A. No.
Q. Do you know whether it used a magnetic compass?
A. I do not know how the -- I do not recall how the position-sensing apparatus worked.
Q. So you don't know whether it had a gyroscope or a differential odometer, for example?
A. That's right.
Q. Or whether it used a steering wheel sensor?
A. That's right.
Q. Isn't it true that the location system, including all the sensors for determining the vehicle position and vehicle heading, was supplied to you from NEC?
A. That is correct.
Q. So you didn't design or implement any part of the location finding hardware for "The Back Seat Driver"; that was obtained from NEC, correct?
A. That is correct.
Q. It's also true that you did not design the speech synthesizer that was used for "The Back Seat Driver"; that was also an off-the-shelf device called DecTalk, correct?

A. That's correct.

Q. That was purchased from where?

A. I don't recall how we obtained it.

Q. But it was an off-the-shelf speech synthesizer, right?

A. At that time one could purchase such an item from Digital Equipment Corporation.

Q. And that's where the DEC in DecTalk comes from?

A. That's correct.

Q. It's Digital Equipment Corporation?

A. That's correct.

Q. Do you know who actually purchased the actual synthesizer that you used for "The Back Seat Driver"?

A. I do not know for sure that it was purchased. It's also possible that it was donated by Digital Equipment Corporation to the media lab.

Q. How did you first come into possession of the DecTalk synthesizer that you used in "The Back Seat Driver"?

A. I don't recall.

Q. And the DecTalk speech synthesizer that was an off-the-shelf speech synthesizer, does that come

Q. When did you reduce to practice the invention that's claimed in claim one of the '685 patent?

MR. BAUER: Objection. Vague. Calls for a legal conclusion.

A. You'd better rephrase that, please.

Q. When did you first -- when were you first reasonably certain that "The Back Seat Driver" would work?

MR. BAUER: Objection.

A. Sometime prior to June of 1989.

Q. How do you know that?

A. Well, by June of 1989, I was already claiming that it was a working system. I don't think that I recall the date in which I first got someone from point A to point B without them getting lost. So June would be a late date, June of 1989 would be a late date.

Q. Did you continue to use "The Back Seat Driver" with other drivers after you knew it would work?

A. Yes.

Q. Did you continue to use "The Back Seat Driver" with other drivers after you knew it would

printed that provide a definition of what you meant by "discourse" when you used that term in the claims; not just examples, but an actual definition of what is and what is not discourse?

MR. BAUER: Objection. Asked and answered. Asking for a legal conclusion.

A. I don't see anything in the thesis that provides a definition of the sort one might use in a dictionary or a textbook for discourse in the patent itself.

Q. Did the "Direction Assistance" program that was running at the Boston Computer Museum, did that generate discourse?

A. The which program, please?

Q. The "Direction Assistance" that was running at the Boston Computer Museum.

A. Yes.

Q. Did the "Direction Assistance" program that was running at the Boston Computer Museum have a discourse generator functionally connected to a computer apparatus?

A. Yes.

Q. Did the discourse generator on the "Direction Assistance" that was on display at the

EXHIBIT

9

(Part 2)

Boston Computer Museum consult a map database?
A. Yes.
Q. Did the discourse generator on the "Direction Assistance" device that was on display at the Boston Computer Museum compose discourse, including instructions and other messages for directing a driver to the destination?
A. Would you repeat that, please.
Q. Did the "Direction Assistance" device that was on display in the Boston Computer Museum include a discourse generator that composed discourse, including instructions and other messages for directing the driver to the destination?
MR. BAUER: Objection. Asking for a legal conclusion. Lack of foundation.
A. I don't think "Direction Assistance" had the other messages aspect.
Q. Did it have everything else?
MR. BAUER: Objection. Everything else meaning what?
MR. LEAVELL: That I just recited in my previous question.
A. Driving instructions. The "Direction Assistance" program was able to process discourse

relating to driving instructions.

Q. Was it able to compose discourse including driving instructions?

A. Yes.

Q. Was it able to direct a driver to a destination?

A. Yes.

Q. Did the "Direction Assistance" that was on display in the Boston Computer Museum have a speech generator?

A. Yes.

Q. Was the speech generator functionally connected to the discourse generator for generating speech from the discourse provided by the discourse generator?

A. Yes.

Q. Did the "Direction Assistance" that was on display in the Boston museum have a voice apparatus functionally connected to the speech generator for communicating said speech provided by said speech generator to a driver?

A. As long as we are to understand a driver to mean the person who wants to get the instructions. They are not, of course, in a car at that time.

you know that the constraint -- well, you probably haven't. The constraints for a museum exhibit are overwhelmingly influenced by two features. One is that there's a lot of kids in the room. They are making a lot of noise. It's hard to hear. The other one is that museum goers are going to stick their fingers anyplace they possibly can, including speaker cones. So you have to protect things from having like Coke spilled on them.

So I speculate, but I feel confident speculating, that the museum personnel were in charge of providing the speaker system so that folks could hear the "Direction Assistance" talk in the din that is the computer museum. Is that responsive?

Q. Yes. Did the "Direction Assistance," as it was installed in the Boston Computer Museum, have a map database that distinguished between physical and legal connectivity?

MR. BAUER: Objection. Undefined terms.

A. You asked did -- please repeat that.

Q. Did the "Direction Assistance," as it was installed in the Boston Computer Museum, have a map database which distinguished between physical and

legal connectivity?
A. Yes, it did.
Q. Did it have that for the entire time it was on display at the computer museum?
A. Yes, it did.
Q. Did that map database have a set of straight line segments and a set of nodes, each end point of which was a pointer to a node representing the coordinates of the end point and a set of other segments which are physically and legally connected to that end point?
A. Yes, it did.
Q. And the "Direction Assistance" had that the entire time it was on display in the computer museum?
A. Yes.
Q. Did the "Direction Assistance" have a map database that was based on DIME files of the United States census extended to represent physical and legal connectivity?
A. Yes, it did.
Q. And it had that the entire time it was on display at the Boston Computer Museum?
A. That is correct.

Q. Did the "Direction Assistance," as displayed at the Boston Computer Museum, have a map database that was a DIME file that was further extended to distinguish bridges, underpasses, tunnels, rotaries and access ramps from other street types?

A. I can't recall.

Q. Did the "Direction Assistance," as it was installed in the Boston Computer Museum, ever have a map database based on TIGER files?

A. I'm not certain.

Q. Did the "Direction Assistance," as it was installed in the Boston computer museum, have a map database that included turn difficulty?

A. I'm not certain.

Q. Is there any way to find out?

A. Yes.

Q. How?

A. We could consult the papers that describe "Direction Assistance," and see whether it had that feature at that time.

Q. Okay. Now, I want to refer to the Exhibit 73.

A. Just a moment, please.

Q. I've handed you Exhibit 73. This is your technical memo No. 1, and this one says that there's a "Direction Assistance" installed at the computer museum in Boston. So is what you're telling me that if this paper describes something as being in "Direction Assistance," then --

MR. BAUER: I'm sorry, I didn't hear the word.

Q. If this paper describes something as being in "Direction Assistance," is it safe to assume, then, that the "Direction Assistance" that was running at the computer museum at the time of this paper had the same features that are described in the paper?

A. Yes, if it, in fact, says that.

Q. Now, on Page 3 of Exhibit 73, the technical paper, there's a paragraph, the third paragraph that begins with "segments in the DIME files."

A. Hmm-hmm.

Q. That's describing physical and legal connectivity, correct?

A. Yes.

Q. Does that accurately describe how the "Direction Assistance" that was running at the

Boston Computer Museum distinguish between physical and legal connectivity?

A. Yes.

Q. And it did that the entire time it was on display at the computer museum, correct?

A. Yes.

Q. The next paragraph in this article talks about streets not being created equal, and each street given a value for goodness. Do you see that paragraph?

A. I do.

Q. Does that accurately describe how the streets were assigned values in the "Direction Assistance" program as it was being run at the Boston Computer Museum?

A. Yes.

Q. And it ran like that the entire time it was on display in the museum? It had that database features the entire time it was on display at the museum, correct?

A. I think that's highly likely.

Q. Is there any doubt?

A. Yes, there's some doubt, because as I think I testified earlier, I'm not absolutely certain that

description of a route finder, and that description carries over to the middle of Page 7. Do you see that?

A. Hmm-hmm.

Q. Does the description of the route finder in this paper correctly describe the route finder that was used by the "Direction Assistance" demonstration that was on display at the Boston Computer Museum in 1987?

A. There is some doubt in my mind.

Q. Why is that?

A. This description talks about best first search, and I do not know at what point A* was introduced. So if this paper is -- let's assume this paper is correct -- rather, let's assume at the time of writing we were using best first search in the "Direction Assistance."

Q. So is it your belief that the "Direction Assistance" as it was installed in the Boston Computer Museum, as it was originally installed there, used a best first search route finder algorithm?

A. Yes.

Q. Did you eventually change that to have an

Q. Did it include gas stations or stores or any types of landmarks?

A. Any types of landmarks?

MR. BAUER: Objection. Vague.

A. Could you be more specific, please?

Q. Did the "Direction Assistance," as it was installed in the Boston Computer Museum in 1987, include a listing of famous places by name?

A. I don't think so.

Q. Did the "Direction Assistance" program that was run at the Boston Computer Museum in 1987, 1988 or early 1989, did it have a system in which the intersection in a route was classified under one type in a taxonomy of intersection types and a discourse generated in relation to each said intersection depended on its type?

A. Now, would you restate the question or repeat the question, please?

Q. Sure. As installed in the Boston Computer Museum in 1987 and 1988, during any of that time, did the "Direction Assistance" have a classification for intersections in a route?

A. With the qualification that I do not recall the date when the "Direction Assistance" was removed

from the museum. So if you'll just restate that with a different date, I can be more confident giving you a fully truthful answer.

Q. As it was installed in 1987, did it have that ability?

A. It had a taxonomy of turns, yes.

Q. And a taxonomy of intersection types?

A. Yes.

Q. And the discourse generated in relation to each intersection depended on its type?

A. That's correct.

Q. And as the "Direction Assistance" was installed in the Boston computer museum, the taxonomy of intersection types included continue, forced turn, U-turn, enter, exit, onto rotary, stay on rotary, exit rotary, fork, turn and stop, correct?

A. Are you reading from the "Direction Assistance" paper?

Q. I read that from claim 43 of the patent.

A. Okay. Without checking more carefully, it's difficult to say. The taxonomy did change over time as I learned better how turns ought to be described. So I could not testify under oath that

C E R T I F I C A T E

I, JAMES R. DAVIS, do hereby certify that I have read the foregoing transcript of my testimony, and further certify that it is a true and accurate record of my testimony (with the exception of the corrections listed below):

Page Line Correction

Signed under the pains and penalties of perjury this day of , 2006.

JAMES R. DAVIS

CORRECTIONS TO DEPOSITION TRANSCRIPT
OF JAMES R. DAVIS
February 16, 2006

MASSACHUSETTS INSTITUTE OF TECHNOLOGY V. HARMAN INTERNATIONAL INDUSTRIES INC., C.A. NO. 05-10990-DPW

Page	Line	Change/Correction	Reason
Global	Global	Change all "media lab" to "Media Lab"	Transcription error
Global	Global	Change "Mr. Pasternak" to "Mr. Pasternack"	Transcription error
29	13-14	Change "general motors" to "General Motors"	Transcription error
38	17	Change "satisfies" to "says"	Transcription error
43	23	Change "a known" to "unknown"	Transcription error
58	18	Change "reasoned" to "recent"	Transcription error
92	3	Change "genius" to "genus"	Transcription error
95	11	Change "ability" to "a built"	Transcription error
102	16	Change "censored" to "censured"	Transcription error
108	14	Change "THE WITNESS" to "MR. LEAVELL"	Transcription error
115	10	Change "sites" to "cites"	Transcription error
115	24	Change "now" to "no"	Transcription error
117	19	Change "rephrase" to "rephrasing"	Transcription error
121	22	Change "band" to "banned"	Transcription error
122	6	Change "band" to "banned"	Transcription error
141	19	Change "never I am imagined" to "never imagined"	Transcription error
143	23	Change "undersatnd" to "understand"	Transcription error
147	8	Change "an affera" to "anaphora"	Transcription error
156	15	Change "no" to "so"	Transcription error
181	9	Change "Pierre Humbert" to "Pierrehumbert"	Transcription error
198	16	Change "limb" to "him"	Transcription error
209	3	Change "sent" to "spent"	Transcription error
245	1	Change "was in the" to "wasn't"	Transcription error

I have read the foregoing transcript of my deposition and except for the corrections and changes noted above, I hereby subscribe to the transcript as an accurate reflection of the statements made by me.

James R. Davis

EXHIBIT

10

Excerpts from the:

February 8, 2006
Deposition of
Christopher M. Schmandt

VOLUME 1

PAGES 1 - 301

EXHIBITS D32 - D44

IN THE UNITED STATES DISTRICT COURT

FOR THE DISTRICT OF MASSACHUSETTS

No. 05-10990 DPW

- -

MASSACHUSETTS INSTITUTE OF TECHNOLOGY,

Plaintiffs

vs.

HARMAN INTERNATIONAL INDUSTRIES, INCORPORATED,

Defendants

- -

VIDEOTAPED DEPOSITION OF CHRISTOPHER M. SCHMANDT

Wednesday, February 8, 2006 9:38 a.m

Proskauer Rose LLP

One International Place, Boston, MA 02111

Reporter: Janet M. Konarski, RMR, CRR

LegaLink Boston

320 Congress Street, Boston, MA 02110

(617)542-0039

APPEARANCES:

PROSKAUER ROSE LLP
(By Steven M. Bauer, Esquire,
and Kimberly A. Mottley, Esquire)*
One International Place
Boston, Massachusetts 02110
(617)526-9616
Counsel for the Plaintiff

KIRKLAND & ELLIS LLP
(By Craig D. Leavell, Esquire)
200 East Randolph Drive
Chicago, Illinois 60601
(312)861-2105
Counsel for the Defendant

ALSO PRESENT:
Robert P. Hart, Chief Intellectual Property
Counsel, Harman International
Jason Lachapelle, Videographer
* Not present at all times

THE VIDEOGRAPHER: Here begins Videotape No. 1 in the deposition of Chris Schmandt in the matter of Massachusetts Institute of Technology v. Harman International Industries, Incorporated in the United States District Court for the District of Massachusetts, Case No. 05-10990DPW. Today's date is February 8, 2006. The time on the video monitor is 9:38 a.m.

The video operator today is Jason Lachapelle, a notary public, contracted by LegaLink Boston. This deposition is taking place at One International Place, Boston, Massachusetts, and was noticed by Kirkland & Ellis for the defense.

Counsel, please voice identify yourself and state whom you represent.

MR. BAUER: Steven Bauer from Proskauer Rose, representing MIT and the witness.

MS. MOTTLEY: Kimberley Mottley from Proskauer Rose, representing MIT and the witness.

MR. LEAVELL: Craig Leavell from Kirkland and Ellis, representing Harman.

MR. HART: Robert Hart representing Harman International.

THE VIDEOGRAPHER: The court reporter

today is Janet Konarski. Would the reporter please swear in the witness.

CHRISTOPHER M. SCHMANDT,

having been duly sworn, after presenting identification in the form of a driver's license, deposes and says as follows:

DIRECT EXAMINATION

BY MR. LEAVELL:

Q. Good morning, sir.

A. Good morning.

Q. We've introduced each other, but for the record, my name is Craig Leavell, and I'll be taking your deposition today. It's important that you understand each question that I ask of you. So, if there is any time that you don't understand a question or any portion of a question that I ask you, will you let me know, so that I can rephrase or try to fix the question?

A. I'll do my best.

Q. It's also important that you hear my questions. If there is a question I ask that you don't hear, will you let me know, so that I can repeat it?

A. Again, I'll do my best.

Q. If at any time today you realize that an

referring to when you say "navigation systems?" Are you referring to car electronics which describe position of the vehicle or are you referring how to give directions electronically?

Q. Why do you draw a distinction between the two?

A. Because they're different. My answer to your question would be different depending on which definition you used.

Q. Considering both of those areas to be within the purview of my question, when did your work begin?

A. I would say 1987, possibly a little bit earlier.

Q. What was the first --

A. Excuse me. Sometime in 1986 or 1987.

Q. And is that your work related to the system that was eventually known as directory assistance?

A. The name of the system was Direction Assistance, and yes.

Q. What was the first thing that you recall doing in connection with Direction Assistance?

A. Discussing it with Jim Davis.

counsel. I'm sorry.

MR. LEAVELL: I got you.

MR. BAUER: Not Fish & Richardson. Patent counsel.

MR. LEAVELL: We don't have to deal with those guys, do we?

MR. BAUER: Mr. Pasternack used to be at Fish a long time ago. That is what I was thinking, but that is just my off-the-record guess, because there were three redacted pages to it.

MR. LEAVELL: Understood.

BY MR. LEAVELL:

Q. Mr. Schmandt, this is the article I was referring to a few minutes ago where it says, and you'll see on the first page, left-hand column, last paragraph, "At the time of this writing, June '89, we have a working system on the road."

A. Mm-hmm.

Q. The video on your website, how does that video compare -- how does the system shown in that video compare to the system that was being --

A. Let me see.

Q. -- that was being demonstrated and working on the road in June of '89?

A. The underlying system, as shown in Figure 2 of the paper, is the same in the paper and in the videotape. The actual behavior of the system, what utterances it generates, I don't know whether the 1990 video would have generated the same utterances in June, 1989, as we were continually, we were continually working on improving the system.

Q. Other than the exact utterances, if there is something that is shown in the video on your website, is it reasonable to assume that whatever we can see in the video on the website was also present in the system that was working on the road in June of '89, or would that be a dangerous assumption?

A. It would be a dangerous assumption, because if you say other than what the system says, that's most of what you "see" when you have a demonstration of a vehicle navigation system.

There is a quick very shot dark -- dark shot into the trunk of a car, which would show a component of the hardware, which would be the same in the two systems. Other than that, most of what you see or hear in this little demo video is what the system says, and that could easily have changed between the two.

having come one that idea?
A. No, I don't.
Q. Did the Direction Assistance device generate discourse?
A. Yes.
Q. Did the discourse generator in the Direction Assistance device consult the map database?
A. Of course.
Q. Why do you say "of course?"
A. Because the discourse generator had the ability to describe different physical topologies, and those physical topologies were stored in a map database.
Q. Did the discourse generator on the Direction Assistance device, was it functionally connected to a computer apparatus?
A. The discourse generator was software running in a computer.
Q. So, that is a yes?
A. I don't know whether that is functionally connected. I think that's inside of.
Q. So, something that is inside of something is not functionally connected to it?
A. It probably is.

Q. With that understanding, was the discourse generator in the Direction Assistance device functionally connected to a computing apparatus?

A. Yes, it was.

Q. Did the discourse generator on the Direction Assistance device compose discourse including instructions and other messages for directing the driver to the destination?

A. It composed instructions. I don't have any recollection of it composing other messages.

Q. In the context of Claim 1 of the '685 patent, which says that, "The discourse generator composes discourse including instructions and other messages for directing the driver to the destination," what does that mean by "other messages"?

A. I think we have quite a few examples in the description. One example could be warnings. You're driving too fast, slow down.

Q. Anything else?

A. There are many other examples in the description.

Q. Like what?

A. Do you want me to find them for you?

Q. Please. I'd like to interrupt you for a

E R R A T A S H E E T

I, CHRISTOPHER M. SCHMANDT, do hereby certify that I have read the foregoing transcript of my testimony, and further certify that it is a true and accurate record of my testimony (with the exception of the corrections listed below).

PAGE LINE CORRECTION

____ ____ ______________________________________

____ ____ ______________________________________

____ ____ ______________________________________

____ ____ ______________________________________

____ ____ ______________________________________

____ ____ ______________________________________

____ ____ ______________________________________

____ ____ ______________________________________

____ ____ ______________________________________

____ ____ ______________________________________

____ ____ ______________________________________

____ ____ ______________________________________

Signed under the pains and penalties this ________ day of ____________________, 2006.

____

CHRISTOPHER M. SCHMANDT

CORRECTIONS TO DEPOSITION TRANSCRIPT OF CHRISTOPHER M. SCHMANDT
February 8, 2006

MASSACHUSETTS INSTITUTE OF TECHNOLOGY V. HARMAN INTERNATIONAL INDUSTRIES INC., C.A. NO. 05-10990-DPW

Page	Line	Change/Correction	Reason
Global	Global	Change "Trobough" to "Trobaugh"	Transcription error
Global	Global	Change "realtime" to "real time"	Transcription error
Global	Global	Change "dime" to "DIME"	Transcription error
10	15	Change "I was" to "They were"	Clarification
51	2	Change "media lab" to "Media Lab"	Transcription error
57	17	Change "you o know" to "you know"	Transcription error
68	20	Change "Stephen Marte" to "Stefan Marti"	Transcription error
69	8	Change "sense" to "since"	Transcription error
78	23	Change "invention" to "information"	Clarification
79	8	Change "invention" to "information"	Clarification
89	4-5	Change "committee on the use of humans in experimental subjects" to "Committee on the Use of Humans as Experimental Subjects"	Transcription error
96	3	Change "Rittbueler" to "Rittmueller"	Transcription error
128	24	Change "now" to "know"	Transcription error
130	8	Change "we're" to "were"	Transcription error
130	9	Change "we're" to "were"	Transcription error
131	21	Change "MS. MOTTLEY" to "MR. HART"	Transcription error
158	1	Change "tie-down" to "tied-down"	Transcription error
182	16	Change "tit" to "it"	Transcription error
186	13	Change "List" to "Lisp"	Transcription error
186	13	Change "Spark" to "Sparc"	Transcription error
202	15	Change "68,000" to "68000"	Transcription error
203	4-6	Change "Yes. 'The computing apparatus appears to be a pair of microprocessors, at least one of which is a Motorola 68,000.'" to "Yes. The computing apparatus appears to be 'a pair of microprocessors, at least one of which is a Motorola 68000.'"	Transcription error
205	8	Change "It's a drawing. So, one of" to "It's a drawing. BY MR. LEAVELL: Q. So, one of"	Transcription error
213	20	Change "their" to "the"	Transcription error
216	15	Change "lane be" to "lane may be"	Transcription error
234	5	Change "list" to "Lisp"	Transcription error
238	11	Change "an affeer of" to "anaphora"	Transcription error
251	6	Change "media library" to "Media Lab Library"	Transcription error

CORRECTIONS TO DEPOSITION TRANSCRIPT
OF CHRISTOPHER M. SCHMANDT
February 8, 2006

MASSACHUSETTS INSTITUTE OF TECHNOLOGY V. HARMAN INTERNATIONAL INDUSTRIES INC., C.A. NO. 05-10990-DPW

Page	Line	Change/Correction	Reason
259	16	Change "degenerative rates" to "degenerates"	Transcription error
278	17	Change "hardware man" to "Harman"	Transcription error
280	12-13	Change "Before Voice Assisted Automobile Navigation" to "Back Seat Driver: voice assisted automobile navigation"	Transcription error
285	10	Change "spark station" to "Sparc Station"	Transcription error
286	4	Change "tiger" to "TIGER"	Transcription error
287	9	Change "deertation" to "dissertation"	Transcription error

I have read the foregoing transcript of my deposition and except for the corrections and changes noted above, I hereby subscribe to the transcript as an accurate reflection of the statements made by me.

Christopher M. Schmandt

EXHIBIT

11

SYNTHETIC SPEECH FOR REAL TIME DIRECTION-GIVING

Christopher M. Schmandt and James Raymond Davis
The Media Laboratory
Massachusetts Institute of Technology

Abstract

The Back Seat Driver is a research prototype of a system to use speech synthesis as a navigational aid for an automobile equipped with localization equipment. We are evaluating the user interface by field trials. As this is work in progress, this paper will primarily give an overview of the system and describe its components. Included will be discussion of the map database, route finding algorithm, repair strategies, and the discourse generator.

With advances in navigation technology and automotive electronics[3,8] has come increasing interest in cars that know where they are and can help you figure out how to reach your destination. Most prototype projects have used various forms of display to present this information, and not all of them have included route finding ability[2,5,7,10,12,13,14,15,18,20,21] For safety reasons, a display may not be particularly suited to this task, moreover there is some evidence that drivers do better following spoken directions than reading maps [19]. Our project, the Back Seat Driver, uses synthetic speech to give driving directions in real time. It plans a route, talks the driver through the route, and not only warns the driver when she has made an error, but also plans an alternate, corrective route.

This paper is an overview describing work in progress. We hope to publish more detailed explanations of the various portions at a later date. At the time of this writing (June, 1989) we have a working system on the road and are simultaneously conducting field trials and improving the direction giving ability and database. Although we do not aspire to prove that voice is better than graphics for direction giving, we do aim to build an optimal system. Early results are very encouraging, suggesting that speech may prove to be a powerful technology in automobiles of the future.

Talking about Directions

There are many factors which contribute to good route description by people, some of which our system only touches on. The problem is complex and simple solutions are not likely to produce comfortable interfaces.

A good route is not simply the shortest, but is more likely to be a combination of the fastest and the easiest to follow. "Easiest to follow" will, however, differ between directions given in advance and directions given in real time by a fellow passenger. Directions given in advance (as e.g. by [4], or the system at Hertz rental counters) must be simple, because the driver alone has the burden of interpreting and following the directions, and there is no help if the driver gets lost. When the direction giver is in the car it is practical to use minor streets or short cuts.

Good directions take into account conceptual *portions* of a route, which make it easier of the driver to keep track of her location on a more global basis. These may include named neighborhoods, types of neighborhoods (business, residential, parks) and types of roads (expressways, parkways, "main" roads, twisty or narrow streets).

By way of example, one of the authors was recently given directions at a car rental counter in a city new to him. The agent at the counter said *"As you leave the airport, keep bearing to the right. You'll go around the end of the runway and see signs for the Interstate north."* The "computerized driving directions" printed at the counter described the same route as 5 separate segments, with mileages and names for each. Especially as it was night, the latter were almost impossible to follow, while the former had succinctly captured the salient aspects of the route.

When the directions are being given by a passenger, the real-time aspect becomes more important. Directions will be given just in time, taking into account vehicle speed, difficulty of the expected maneuver, driving styles, and road, weather, and traffic conditions. During long highway stretches with little need for description, the direction giver must maintain the driver's confidence. The passenger will also be watching for errors and trying to warn against them, again based on fine observations of the vehicle's speed and

Manuscript received June 9, 1989

0098 3063/89/0200 0649$01.00

direction. When a mistake is made, the passenger will tell the driver about it and together they will take corrective action (which is unlikely to be simply a sudden stop!).

Project Goals

The main goal of this project is to evaluate the utility of speech synthesis as the user interface to a real-time navigation system in an urban environment. Of particular concern is the discourse structure:

- how should driving acts be described?
- how can a description be generated from a route?
- how should timing considerations be applied?
- what kinds of feedback, both positive and negative, does the user require?
- what kinds of visual cues are most useful in describing an approaching location?

This information is gained from both laboratory simulations and field trials.

Our desire is to build the best possible real-time route describer. Although we believe a speech interface to the navigation unit is superior and safer than a visual interface, we do not plan to conduct direct comparison studies.

In the course of field trials to evaluate and improve our automatic direction giving, we hope to specify key components of the map database. We expect discourse behavior may need to vary with conditions (traffic, weather, day/night). It is likely that different visual cues may be useful in these situations. All must be represented in the database.

Geographic Database

Our database covers 41 square miles in the Boston area, including parts of Boston, Cambridge, Brookline, Somerville, and Watertown. It originated as a DIME (Dual Independent Map Encoding) file distributed by the United States Geological Survey[1]. A DIME file consists of a set of straight line segments, each with a name, a type, endpoints in longitude and latitude, and some additional information such as address numbers. Initially our database contained many errors. Correcting them required actually traveling most of the segments.

A DIME file alone is not sufficient for finding routes. The DIME files indicate physical connectivity, but route finding requires *legal* connectivity, i.e., one can legally drive from one segment to the next (one way streets are a simple example). We extended the data base format to explicitly represent legal connectivity. Since some streets are better than others, the database must include a measures of *quality*. We take this to be a largely subjective measure of the ease of locating and following a street. This allows the route finder to prefer to use streets of higher quality.

The generation of easily followed natural descriptions requires more extensions. We added a number of new segment types to distinguish bridges, underpasses, tunnels, rotaries, and access ramps. All these extensions were done for an earlier route finding project[4].

We are presently adding landmarks to the database. Drivers need landmarks to know how far to drive and when to turn. If the Back Seat Driver had eyes, it could simply choose landmarks as needed by looking for them in the landscape. Being blind, it must rely on landmarks coded into the map database. We have added traffic lights, stop signs, and some buildings to the portions of the landmark database. A main task now is to determine what else must be added.

In addition to landmarks, other information is useful for providing assistance following a route. We found it very useful to add lane information, both number of lanes as well as any turn restrictions on lanes (e.g. left turn only). On short street segments, it is important to give lane advice (*"After the turn you'll want to get into the left hand lane."*) or else the driver may be unable to make the following turn. Lane warnings (*"Stay out of the left turn lane."*) are also important driving cues.

An interesting problem arises at complex intersections, typically a maze of ramps between major arteries, possibly at different elevations (see figure 1). Such intersections are typically not accurately recorded in the map. Furthermore, limitations in the resolution of the position tracking equipment make it difficult to distinguish one segment from another, especially as they are likely to diverge at narrow angles. The combination of uncertainties in the map and uncertainties in position make it difficult to give a clear spoken directions. Fortunately these intersections are usually well signed, so the Back Seat Driver can give directions by referring to the signs, e.g. *"Follow the signs to the expressways and airport"*. The wording of these signs needs to be in the database. It is important that Back Seat Driver's understand what the sign says, not simply utter the words. There are two reasons for this. First, our internal representation for text is based on syntactic structure, not text strings. Second, the objects mentioned in the signs (cities and roads) should be entered into the discourse model. They should become salient for future reference. This means that the text of a sign must be parsed, so that e.g. the sign text "Cambridge, Somerville, and Storrow Drive" should become a conjunction of the two cities "Cambridge" and "Somerville" and the street named "Storrow Drive".

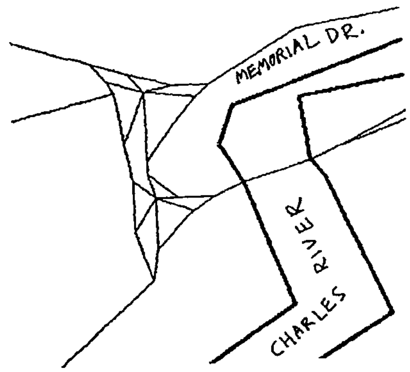


Figure 1: Access ramps at an interchange

System

Our vehicle is equipped with a localization unit built by NEC Home Electronics, Ltd., the project sponsor. It is a dead-reckoning position keeping system which uses speed and direction sensors. To compensate for error, it uses map matching on a map database stored on CD ROM. The system described more fully in [16,17]

As this is a research prototype, much of the computation is done in a base station computer laboratory (on a Symbolics Lisp Machine), rather than a computer on the vehicle. Two cellular telephones link the computer to the car. The on board navigational hardware transmits position and velocity via modem and cellular phone to the base station. The base station computer does all route planning and discourse generation. Speech synthesis is performed in a commercial text-to-speech synthesizer (Dectalk) cabled to the Lisp Machine. Synthesized instructions to the driver are relayed via the second cellular link and a speaker phone in the car. The keypad of the second phone also serves as the driver's control unit for the Back Seat Driver. Through this keypad a driver selects a destination, requests repeats of spoken information, and accesses other services of the Back Seat Driver.

A block diagram of the system appears in figure 2.

Routes

The Back Seat Driver can find the shortest, fastest, or most easily followed route. Route finding uses an A* search algorithm[11]. Depending on the driver's preference, one of three cost metrics is used, The *distance* metric is simply the sum of the lengths of the segments which compose a route. The *speed* metric scales each distance by a factor dependent on street quality. In addition, penalties are incurred for turns (more for left turns than right), stop signs, and traffic lights (which may be red). The *simplicity* metric, following [6], seeks to minimize the number of turns by imposing a distance penalty for each turn.

Discourse Strategies

The instructions are detailed and natural, and include a rich taxonomy of driving verbs. The dialog system uses cues such as vehicle speed and difficulty of driving actions to attempt to deliver instructions at the proper pace and in a timely manner. In addition, the system can anticipate some of the driver's possible mistakes and give warnings to avoid them.

Describing a route requires going from a series of segments (typically city blocks) in the database to a series of travel segments which will be separated by decision points. For example, going straight down a main street for five blocks will not be thought of by the driver as five separate acts, but rather one street traversal. A key piece of this analysis is that the need to make a decision is based on knowledge of what is *obvious*. Drivers do not want to be nagged at each corner to continue straight, but when they come to a questionable fork in the road they do want to be told which way to proceed.

If the driver does make a wrong turn, or misses a turn, the Back Seat Driver describes the error and then incrementally calculates a new route, rather than simply backtracking to the point of the error. Route planning includes weighting for length of the trip, difficulty of driving maneuvers (such as left turns against traffic), street quality, and complexity of the spoken directions.

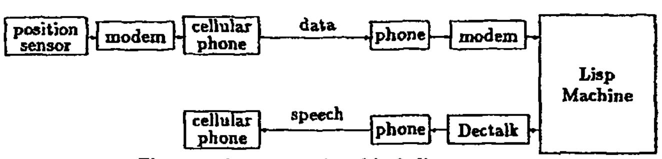


Figure 2: Communications block diagram

MIT 00935

As opposed to much prior work in discourse generation, the Back Seat Driver is a real-time system which must factor in a number of temporal considerations. It needs to give each stage in the directions at just the right point, in terms of the time it takes to execute the driving maneuver as well as the speed of the vehicle approaching the intersection. For safety considerations, we would rather err on the side of giving the driver plenty of warning, but a cue given too far in advance may be miscued (e.g., a turn taken at an earlier intersection). Additionally, the software must consider the length of time it will take to recite an utterance. It is better to miss a turn and plan a new route than start describing the turn at a time when it may be unsafe to execute it (i.e., already well into an intersection).

There are several reasons to give instructions before the act, if time permits. One is to allow the driver to hear the instructions several times, and the other is to allow time to prepare for some acts, e.g., turns from a multi-lane street. These advance notices are lower priority than the description of the act itself, according to an internal set of system goals. Thus, they can be presented if there is adequate time, but will be ignored if the vehicle is approaching the next decision point too quickly.

Reassuring

While the driver is following a route, the system adopts a persistent goal of keeping the user reassured about her progress and the system's reliability. If Back Seat Driver were a human, this might be unnecessary, since the driver could see for herself whether the navigator was awake and attending to the road and driver. But the driver can not see the system, and so needs some periodic evidence that the system is still there.

One piece of evidence is the safety warnings the system gives (e.g. *"slow down"* before a turn), but if all is going well, there will not be any. The system gives two other kinds of evidence that things are going well. First, when the user completes an action, the system acknowledges the driver's correct action, saying something like *"nice work"* or *"good"*. This feature is very popular with most test drivers.

The second form of evidence is to make insignificant remarks about the roads nearby, the weather, and so on. If the driver assumes that the navigator is being cooperative, as set out in Grice's maxims of cooperative conversation [9], then the driver can infer that everything is going well, for otherwise the navigator would not speak of trivial matters. It isn't clear, however, that one really wants a chatty speech synthesizer. Certainly this feature could be useful in a rented car in a new city, where it might actually have some interesting things to say.

Summary

The following is one of the more complex utterances of the Back Seat Driver to date. It summarizes many key points mentioned in this paper, and indicates the current state of operability of the discourse generator:

> Get in the left lane because you're going to take a left at the next set of lights. It's a complicated intersection because there are two streets on the left. You want the sharper of the two. It's also the better of them. After the turn, get into the right lane.

The Back Seat Driver is already working in prototype form. Our present concerns are to determine what a spoken driving assistant should say, to understand how time and speed affect this decision, and to learn what features a map database must have to support generation of instructions.

Acknowledgments

The authors wish to gratefully acknowledge the support of NEC Home Electronics, Ltd.

References

[1] *Geographic Base File GBDF/DIME: 1980 Technical Documentation.* U.S. Department of Commerce, Data Users Services Division, 1980.

[2] Peter Braegas. Function, Equipment, and Field Testing of a Route Guidance and Information System for Drivers (ALI). *IEEE Transactions on Vehicular Technology*, VT-29(2):216–225, May 1980.

[3] Donald F. Cooke. Vehicle Navigation Appliances. In *AUTO CARTO 7 International Symposium on Automation in Cartography*, pages 108–115, March 1985.

[4] James R. Davis and Thomas F. Trobaugh. *Direction Assistance.* Technical Report 1, MIT Media Laboratory Speech Group, Dec 1987.

[5] Ronald A. Dork. Satellite Navigation Systems for Land Vehicles. In *IEEE Position and Location Symposium*, pages 414–418, 1986. IEEE 86CH2365-5.

[6] R. J. Elliot and M. E. Lesk. Route finding in street maps by computers and people. In *Proceedings of the National Conference on Artificial Intelligence*, pages 258–261, 1982.

[7] Edward J. Krakiwsky et al. Research into electronic maps and automatic vehicle location. In *AUTO CARTO 8 International Symposium on Automation in Cartography*, 1987.

[8] Robert L. French. Automobile Navigation: Where is it going? In *IEEE Position and Location Symposium*, pages 406–413, 1986. IEEE 86CH2365-5.

[9] H. P. Grice. Logic and conversation. In Cole and Morgan, editors, *Syntax and Semantics: Speech Acts*, pages 41–58, Academic Press, 1975.

[10] Peter Haeussermann. *On Board Computer System for Navigation, Orientation, and Route Optimization.* Technical Paper Series 840483, Society of Automotive Engineers, 1984.

[11] P. E. Hart, N. J. Nilsson, and B. Raphael. A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on SSC*, 4:100–107, 1968.

[12] Stanley K. Honey, Marvin S. White Jr., and Walter B. Zavoli. Extending Low Cost Land Navigation Into Systems Information Distribution and Control. In *IEEE Position and Location Symposium*, pages 439–444, 1986. IEEE 86CH2365-5.

[13] Toshiyuki Itoh, Yasuhiko Okada, Akira Endoh, and Kenji Suzuki. *Navigation Systems Using GPS for Vehicles.* Technical Paper Series 861360, Society of Automotive Engineers, 1986.

[14] M. D. Kotzin and A. P. van den Heuvel. Dead Reckoning Vehicle Location using a solid state rate gyro. In *Proceedings of the 31rst IEEE Vehicular Technology Conference*, pages 169–172, April 1981. IEEE publication 81CH1638-6.

[15] Ernst Peter Neukirchner and Wolf Zechnall. Digital Map Data Bases for Autonomous Vehicle Navigation Systems. In *IEEE Position and Location Symposium*, pages 320–324, 1986. IEEE 86CH2365-5.

[16] Osamu Ono, Hidemi Ooe, and Masahiro Sakamoto. CD-ROM Assisted Navigation System. In *Digest of Technical Papers*, pages 118–119, IEEE ICCE, 1988.

[17] Osamu Ono, Hidemi Ooe, and Masahiro Sakamoto. Navigation and communication system. In *Digest of Technical Papers*, IEEE ICCE, 1989.

[18] Otmar Pilsak. *EVA: An electronic Traffic Pilot for Motorists.* Technical Paper Series 860346, Society of Automotive Engineers, 1986.

[19] Lynn A. Streeter, Diane Vitello, and Susan A. Wonsiewicz. How to tell people where to go: comparing navigational aids. *International Journal of Man/Machine Systems*, 22(5):549–562, May 1985.

[20] Katsutoshi Tagami et al. *Electro Gyro-Cator: New Intertial Navigation System for Use in Automobiles.* Technical Paper Series 830659, Society of Automotive Engineers, 1983.

[21] Walter B. Zavoli and Stanley K. Honey. Map Matching Augmented Dead Reckoning. In *Proceedings of the 35th IEEE Vehicular Technology Conference*, pages 359–362, 1986. IEEE CH2308-5.

MR. SCHMANDT received his B.S. in Computer Science from MIT and an M.S. in computer graphics from MIT's Architecture Machine Group. He is currently a Principal Research Scientist and director of the Speech Research Group of the Media Laboratory at M.I.T.

His research interests are focused on interactive computer systems and human-interface issues of synchronous and asynchronous communication. His work emphasizes voice interaction for telecommunication based applications, with a goal of describing and then emulating human conversational behavior.

MIT 00937

EXHIBIT

12

FILED UNDER SEAL

EXHIBIT

13

FILED UNDER SEAL

EXHIBIT

14

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF MASSACHUSETTS

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, **Plaintiff,** **v.** **HARMAN INTERNATIONAL INDUSTRIES, INCORPORATED,** **Defendant.**	**Case No: 05-10990 DPW** **Hon. Douglas P. Woodlock**

MIT'S RESPONSES TO HARMAN'S SECOND SET OF REQUESTS
FOR THE PRODUCTION OF DOCUMENTS AND THINGS (Nos. 30-61)

Pursuant to Rules 26 and 34 of the Federal Rules of Civil Procedure, Plaintiff, Massachusetts Institute of Technology ("MIT") submits the following responses and objections to Harman International Industries, Incorporated's ("Harman's") Second Set of Requests for the Production of Documents and Things (Nos. 27-58) (the "Requests")[1]:

GENERAL OBJECTIONS

The following general statements and objections are incorporated into each of MIT's responses, as set forth there in full, even if not repeated therein:

1. MIT objects to the Requests to the extent that they seek documents not relevant to the subject matter of the present lawsuit and/or are not reasonably calculated to lead to the discovery of admissible evidence.

[1] Although Harman identified its First Set of Requests as including "Nos. 1-28" under the caption and in the Certificate of Service, Harman actually served twenty-nine Requests, all of which MIT responded to or objected to. Harman's Second Set of Requests incorrectly begins with number 27. Harman's Second Set of Requests should be numbered 30-61. MIT's responses herein correspond to the renumbered requests.

2. MIT objects to the Requests to the extent that they seek documents or materials that constitute or contain information protected by the attorney-client privilege, the work-product doctrine, and/or any other applicable privilege. Any disclosure of any document that is subject to such a privilege or protection is inadvertent and shall not constitute a waiver of any such privilege or protection, and any such document should, if located by Harman, be promptly returned to MIT. Any information or documents withheld on these grounds will be identified on a separate privilege log, except for such documents that were created or that came into being after initiation of this lawsuit.

3. MIT objects generally to the Requests to the extent they seek confidential documents or materials. MIT will produce confidential information only subject to the protective order entered by the Court that limits the use of documents and information produced by MIT in this action to this proceeding and protects any documents containing MIT confidential information called for by Harman's Requests from unauthorized disclosure to Harman representatives or any third parties.

4. MIT objects generally to the Requests to the extent that they attempt or purport to impose obligations on MIT beyond those required by the Federal Rules of Civil Procedure 26 and 34, or the Local Rules of Practice of the United States District Court for the District of Massachusetts.

5. MIT objects to the Requests to the extent that they seek documents already in Harman's possession, equally accessible to Harman and/or publicly available.

6. MIT objects generally to the Requests to the extent that they seek documents or materials not in MIT's possession, custody, or control.

7. MIT objects generally to the Requests to the extent that they seek to characterize evidence in this matter. To the extent that MIT expressly or impliedly adopts any term used by the Requests, such adoption is specifically limited to these responses and does not constitute an admission of fact or law in this action or any other matter.

8. None of the General Objections and Responses herein is a direct or indirect admission (i) of the truth or accuracy of any statement or characterization asserted by Harman in any pleading or other filing with the Court; (ii) of the validity of any objection or response by Harman to any discovery Request propounded in this action by MIT; or (iii) that any discovery Request propounded by MIT in this action is wholly or partially objectionable under any applicable law or rules.

9. MIT generally objects to the Requests to the extent that MIT has not completed its investigation of the facts in this case and has not completed discovery in this action. Any responses to these Requests are based on information presently known to MIT and are without prejudice to MIT's right to supplement its responses and produce as evidence, any subsequently discovered documents.

10. MIT generally objects to the Requests to the extent they are overly broad or unduly burdensome, including without limitation Requests that require production of "all" or "any" documents.

11. MIT generally objects to the Requests to the extent they are vague, ambiguous, confusing, incomprehensible and/or unanswerable because of undefined or ill-defined terms and/or confusing syntax, or they fail to describe with reasonable particularity the documents sought.

12. MIT's willingness to provide documents or things is not a concession that the subject matter of the particular document is discoverable, relevant to this action, or admissible as evidence.

13. MIT expressly reserves all objections as to relevance, authenticity, and/or admissibility of any documents produced.

14. MIT has responded to the Requests as it interprets and understands each Request made therein. If Harman subsequently asserts an interpretation of any Request that differs from the understanding of MIT, MIT reserves the right to supplement its objections and responses.

15. The presence or absence of any general or specific objection does not mean that MIT does not object on any other grounds.

16. MIT incorporates its General Objections to Harman's First Set of Requests for the Production of Documents and Things (Nos. 1-28 [renumbered 1-29]) as if fully set forth herein.

17. MIT incorporates its General Objections to Harman's First Set of Interrogatories (Nos. 1-7) and Harman's Second Set of Interrogatories (Nos. 8-20) as if fully set forth herein.

18. MIT objects to the production of any document falling within the General Objections set forth above. A statement by MIT that it will produce responsive, non-privileged, and non-work product protected documents means that MIT will produce documents, if any exist, within its possession, custody, or control that (a) are responsive to the particular Request or fall within the particular description in question, (b) have been located by MIT after a reasonable search, and (c) do not fall within one of the General Objections set forth above or within any of the objections contained in the specific responses set forth below.

19. MIT generally objects to the Requests to the extent that they are duplicative of or call for documents and things that MIT has already produced responsive to Harman's First Set of

Requests for the Production of Documents and Things (Nos. 1-28 [renumbered 1-29]) or Harman's First Set of Interrogatories (Nos. 1-7).

20. Subject to these General Objections and the Specific Objections set forth below, MIT will produce responsive documents at a time and place mutually agreed upon by the parties.

SPECIFIC OBJECTIONS AND RESPONSES

REQUEST NO. 27 [RENUMBERED 30]

All documents and things referring or relating to the work of Martin Thoone, Thoone or the Phillips CARIN system, including but not limited to, all reports or correspondence provided to MIT, James R. Davis, or Christopher Schmandt.

RESPONSE TO REQUEST NO. 27 [30]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT and are equally available to Harman. MIT further objects to this Request as duplicative of Request Nos. 12, 13, 14, 17, and 18 from Harman's First Set of Requests for the Production of Documents and Things.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 28 [RENUMBERED 31]

All documents and things referring or relating to Navigation Technologies, Inc. (NavTech), NavTeq or Karlin & Collins, Inc., or any of their systems, such as Driver Guide,

ROGUE, etc., including but not limited to all reports or correspondence provided to MIT, James R. Davis or Christopher Schmandt.

RESPONSE TO REQUEST NO. 28 [31]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT and are equally available to Harman. MIT further objects to this Request as duplicative of Request Nos. 13, 14, 17, and 18 from Harman's First Set of Requests for the Production of Documents and Things.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 29 [RENUMBERED 32]

All documents and things referring or relating to the work of Otmar Pilsak or the Bosch-Blaupunkt EVA system, including but not limited to all reports or correspondence provided to MIT, James R. Davis or Christopher Schmandt.

RESPONSE TO REQUEST NO. 29 [32]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT and are equally available to Harman. MIT further

objects to this Request as duplicative of Request Nos. 1, 6, 12, 13, 14, 17, and 18 from Harman's First Set of Requests for the Production of Documents and Things.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 30 [RENUMBERED 33]

All documents and things referring or relating to the Toyota ElectroMultivision navigation device, including but not limited to all reports or correspondence provided to MIT, James R. Davis or Christopher Schmandt.

RESPONSE TO REQUEST NO. 30 [33]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT and are equally available to Harman. MIT further objects to this Request as duplicative of Request Nos. 12, 13, 14, 17, and 18 from Harman's First Set of Requests for the Production of Documents and Things.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 31 [RENUMBERED 34]

All documents and things referring or relating to the work of Davis and Trobaugh, including but not limited to all documents and things referring to their article "Direction Assistance."

RESPONSE TO REQUEST NO. 31[34]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT and are equally available to Harman. MIT further objects to this Request as duplicative of Request Nos. 1, 6, 7, 8, and 21 from Harman's First Set of Requests for the Production of Documents and Things.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 32 [RENUMBERED 35]

All documents and things referring or relating to the public use of the Direction Assistance display in the Computer Museum in Boston or elsewhere.

RESPONSE TO REQUEST NO. 32 [35]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request to the extent it seeks documents that

are beyond the custody or control of MIT and are equally available to Harman. MIT further objects to this Request as premature to the extent it calls for a legal conclusion as to public use. MIT further objects to this Request as duplicative of Request Nos. 1, 6, 7, 8, and 21 from Harman's First Set of Requests for the Production of Documents and Things.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 33 [RENUMBERED 36]

All documents and things referring or relating to field trials or other demonstrations of the Back Seat Driver system.

RESPONSE TO REQUEST NO. 33 [36]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT and are equally available to Harman. MIT further objects to this Request as duplicative of Request Nos. 9, 10, and 11 from Harman's First Set of Requests for the Production of Documents and Things.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 34 [RENUMBERED 37]

All documents and things considered during the prosecution of the '685 patent, including, but not limited to each reference identified in the September 1990 Information Disclosure Statement.

RESPONSE TO REQUEST NO. 34 [37]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request as duplicative of Request Nos. 1, 6, 16, and 17 from Harman's First Set of Requests for the Production of Documents and Things.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 35 [RENUMBERED 38]

All documents and things referring or relating to "discourse", "discourse generation" or "discourse generators", including but not limited to presentation materials, papers, articles, correspondence, and notes.

RESPONSE TO REQUEST NO. 35 [38]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request as premature to the extent

it purports to characterize or otherwise construe the claims of the '685 patent. MIT further objects to this Request as premature to the extent it calls for a legal conclusion with respect to construction of the claims of the '685 patent. MIT further objects to this Request as not limited to a particular period of time. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity.

Subject to and without waiving the foregoing general and specific objections and in addition to responsive documents that have already been produced, MIT will provide Harman with access to publicly-available materials in MIT's library for inspection of documents responsive to this Request at an agreed-upon time. MIT further identifies as responsive to this Request the following list of materials from Mr. Schmandt's files, which MIT will make available for inspection at an agreed-upon time and place.

- Computational Models of Discourse. Brady and Berwick (eds.). (Cambridge, MA: MIT Press, 1984 (second printing));
- Readings in Natural Language Processing. Grosz, Sparck Jones, and Webber (eds.). (Morgan Kaufman, 1986);
- Pragmatics. Stephen C. Levinson. (Cambridge: Cambridge University Press, 1983);
- Text generation - using discourse strategies and focus constraints to generate natural language text. Kathleen McKeown. (Cambridge: Cambridge University Press, 1985); and
- Computational linguistics - an introduction. Ralph Grishman. (Cambridge: Cambridge University Press, 1986).

- Intentions In Communication, Cohen, Morgan & Pollack (eds.), MIT Press Cambridge, Massachusetts (1990). Pierrehumbert, Julia & Hirschberg, Julia "The Meaning of Intonational Contours In The Interpretation Of Discourse" (Chapter 14).

- Lieberman, Phillip. *Intonation, Perception & Language* (Cambridge, MA: MIT Press, 1967). Chapter 7, "Prominence, Stress, and Emphasis in American English".

- Liberman, Mark & Prince, Alan. On Stress and Linguistic Rhythm. *Linguistic Inquiry,* 8:2 (Spring 1977), 249-336.

- Pierrehumbert, Julia & Hirschberg, Julia. The Intonational Structuring Of Discourse. *Association For Computational Linguistics*, 1985 Proceedings.

- Pierrehumbert, Julia & Hirschberg, Julia. The Intonational Structuring Of Discourse. *Association For Computational Linguistics*, 1986 Proceedings.

- Hirschberg, Julia & Litman, Diane. Now Let's Talk About Now: Identifying Cue Phrases Intonationally. *Association For Computational Linguistics*, 1987 Proceedings.

- Grosz, Barbara & Sidner, Candace. Attention, Intentions, And The Structure Of Discourse. *Computational Linguistics*, 12:3 (July-September 1986), 175-204.

- Grosz, Barbara & Hisrschberg, Julia. Some Intonational Characteristics of Discourse Structure. *Proceedings of ICSLP-92*, I, 429-432.

- Frick, Robert. Communicating Emotion: The Role of Prosodic Features. *Psychological Bulletin*, 9:3(1985), 412-429.

- Duncan, Jr. Starkey. Some Signals And Rules For Taking Speaking Turns In Conversations. *Journal Of Personality & Social Psychology*, 23:2 (1972), 283-292.

- Cruttenden. *Intonation.* (Cambridge: Cambridge University Press, 1986). "The Functions Of Intonation" (Chapter 4).

- Clark, Herbert. *Arenas of Language Use.* (The University Of Chicago Press & Center For The Study Of Language & Information, 1992). "Definite Reference And Mutual Knowledge" (Chapter 1, with Catherine R. Marshall).

- Brown, Gillian; Currie, Karen & Kenworthy, Joanne. *Questions of Intonation* (Taylor & Francis Books, Ltd., 1980). Chapter 2, 21-39.

REQUEST NO. 36 [RENUMBERED 39]

All documents and things referring or relating to "physical connectivity", "legal connectivity" or a "computing apparatus which distinguishes between physical and legal connectivity."

RESPONSE TO REQUEST NO. 36 [39]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request as premature to the extent it purports to characterize or otherwise construe the claims of the '685 patent. MIT further objects to this Request as premature to the extent it calls for a legal conclusion with respect to construction of the claims of the '685 patent. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 37 [RENUMBERED 40]

All documents and things referring or relating to efforts by MIT to enforce the '685 patent.

RESPONSE TO REQUEST NO. 37 [40]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request as duplicative of Request Nos. 3, 4,

5, 12, 13, 14, and 15 from Harman's First Set of Requests for the Production of Documents and Things.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 38 [RENUMBERED 41]

All documents and things referring or relating to MIT's evaluation of a reasonable royalty as a result of any alleged infringement by Harman of the '685 patent, including but not limited to all documents and things referring or relating to reasonable royalty factors as laid out in Georgia Pacifica Corp. v. United States Plywood Corp., 318 F. Supp. 116 (S.D.N.Y. 1970).

RESPONSE TO REQUEST NO. 38 [41]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 39 [RENUMBERED 42]

All documents referring or relating to MIT's alleged proper measure of royalties for any infringement of the '685 patent, including any license agreements.

RESPONSE TO REQUEST NO. 39 [42]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 40 [RENUMBERED 43]

All documents and things upon which MIT intends to rely in support of any "damages" that MIT claims as a result of any alleged infringement by Harman of the MIT patents.

RESPONSE TO REQUEST NO. 40 [43]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 41 [RENUMBERED 44]

All documents and things that MIT intends to offer into evidence in this litigation.

RESPONSE TO REQUEST NO. 41 [44]

MIT objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request as premature as discovery is ongoing. MIT will identify its trial exhibit list when required by the Court (currently set for October 6, 2006).

REQUEST NO. 42 [RENUMBERED 45]

All documents that MIT intends to rely on within the context of any hearing or trial of this matter, including any hearing that may take place with respect to claim construction.

RESPONSE TO REQUEST NO. 42 [45]

MIT objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request as premature as discovery is ongoing. MIT will identify its documents it intends to rely on at any hearing or trial when required by the Court.

REQUEST NO. 43 [RENUMBERED 46]

All documents referred to or identified in MIT's responses to Harman's First or Second Sets of Interrogatories, all documents which those interrogatories Request to be identified, all documents relied upon or consulted in formulating MIT's responses to those interrogatories, and all documents otherwise referring or relating to MIT's responses to those interrogatories.

RESPONSE TO REQUEST NO. 43 [46]

MIT objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 44 [RENUMBERED 47]

All documents provided by MIT (or its counsel) to any expert witness(es) or consultant(s).

RESPONSE TO REQUEST NO. 44 [47]

MIT objects to this Request as vague and ambiguous to the extent that it is not limited to the present litigation. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request as premature as expert discovery is ongoing.

REQUEST NO. 45 [RENUMBERED 48]

All documents provided to MIT (or its counsel) by any expert witness(es) or consultant(s).

RESPONSE TO REQUEST NO. 45 [48]

See Response to Request 44 [47].

REQUEST NO. 46 [RENUMBERED 49]

All bills, statements or invoices provided to MIT (or its counsel) by any expert witness or consultants.

RESPONSE TO REQUEST NO. 46 [49]

See Response to Request 44 [47].

REQUEST NO. 47 [RENUMBERED 50]

All documents and things referring to or evidencing prototypes or demonstrations of the Back Seat Driver invention.

RESPONSE TO REQUEST NO. 47 [50]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available. MIT further objects to this Request as duplicative of Request Nos. 1, 6, 7, 8, 9, 10, and 11 from Harman's First Set of Requests for the Production of Documents and Things and Interrogatories Nos. 6 & 7 from Harman's First Set of Interrogatories (Nos. 1-7).

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 48 [RENUMBERED 51]

All communications between MIT, James R. Davis or Christopher Schmandt and third parties referring or relating to this litigation.

RESPONSE TO REQUEST NO. 48 [51]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available. MIT further objects to this Request as duplicative of Request Nos. 7 and 15 from Harman's First Set of Requests for the Production of Documents and Things.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 49 [RENUMBERED 52]

All documents and things that mention, concern or discuss or are in any way related to the priority date for any claim of the '685 patent.

RESPONSE TO REQUEST NO. 49 [52]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request to the extent it seeks documents that

are beyond the custody or control of MIT. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available. MIT further objects to this Request as vague or ambiguous to the extent that the Request does not specify a time period or subject matter.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 50 [RENUMBERED 53]

All documents and things that support, refer or relate to, MIT's contention that MIT's patents are valid or not invalid.

RESPONSE TO REQUEST NO. 50 [53]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request as premature to the extent that it calls for legal conclusions with respect to validity. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 51 [RENUMBERED 54]

All documents and things supporting a finding of obviousness or non-obviousness of the '685 patent, including any documents showing any nexus between any secondary considerations of non-obviousness and the subject matters claimed in the '685 patent.

RESPONSE TO REQUEST NO. 51 [54]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request as premature to the extent that it calls for legal conclusions with respect to obviousness, non-obviousness, and/or secondary considerations. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 52 [RENUMBERED 55]

All documents and things to support any alleged diligence by, or on behalf of, Mr. Davis and/or Mr. Schmandt in reducing to practice the subject matter recited in the '685 patent.

RESPONSE TO REQUEST NO. 52 [55]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks

EXHIBIT

14

(Part 2)

documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request as premature to the extent that it calls for legal conclusions with respect to diligence and/or reduction to practice. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 53 [RENUMBERED 56]

All documents and things that support, refer or relate to MIT's contention that the '685 patent is enforceable or not unenforceable.

RESPONSE TO REQUEST NO. 53 [56]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request as premature to the extent that it calls for legal conclusions with respect to enforceability of the '685 patent. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 54 [RENUMBERED 57]

All documents and things that support, or refer or relate to, MIT's contention that Harman's products and/or process allegedly infringe the '685 patent.

RESPONSE TO REQUEST NO. 54 [57]

MIT objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request as premature to the extent that it calls for legal conclusions with respect to infringement. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 55 [RENUMBERED 58]

All documents and things that support, or refer or relate to, MIT's contention that Harman willfully infringed the '685 patent.

RESPONSE TO REQUEST NO. 55 [58]

MIT objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request as premature to the extent that it calls for legal conclusions with respect to willful infringement. MIT further objects to this Request to the extent it seeks documents that are beyond the custody or control of MIT. MIT further objects to this Request to the extent that any

responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 56 [RENUMBERED 59]

All documents and things that support, or refer or relate to, MIT's contention that MIT is entitled to damages for Harman's alleged infringement of the '685 patent.

RESPONSE TO REQUEST NO. 56 [59]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 57 [RENUMBERED 60]

All documents and things having any tendency to prove or disprove each and every affirmative defense asserted by MIT in its responsive pleadings to Harman's counterclaims.

RESPONSE TO REQUEST NO. 57 [60]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity. MIT further objects to this Request to the extent that any responsive documents are already in Harman's possession, readily accessible to Harman and/or publicly available.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

REQUEST NO. 58 [RENUMBERED 61]

All documents and things having any tendency to prove or disprove each and every allegation not admitted by MIT in its responsive pleadings to Harman's Counterclaims.

RESPONSE TO REQUEST NO. 58 [61]

MIT objects to this Request as overbroad, unduly burdensome, and unlikely to lead to the discovery of admissible evidence. MIT further objects to this Request to the extent it seeks documents protected by attorney-client privilege, the work-product doctrine or any other privilege or immunity.

Subject to and without waiving the foregoing general and specific objections, MIT states that it has already produced or logged on its privilege log any documents in its possession, custody, or control responsive to this Request.

Dated: April 21, 2006

Respectfully submitted,

Massachusetts Institute of Technology,

By its Attorneys,

/s/ John W. Pint

Steven M. Bauer (BBO# 542531)
Kimberly A. Mottley (BBO# 651190)
John W. Pint (BBO# 660548)
PROSKAUER ROSE LLP
One International Place
Boston, Massachusetts 02110-2600
Phone: 617-526-9600
Fax: 617-526-9899

CERTIFICATE OF SERVICE

I HEREBY CERTIFY that on April 21, 2006, I caused a true and correct copy of MIT'S RESPONSES TO HARMAN'S SECOND SET OF REQUESTS FOR THE PRODUCTION OF DOCUMENTS AND THINGS (NOS. 30-61) to be served on the following counsel of record via email:

Robert J. Muldoon, Jr.
SHERIN AND LODGEN, LLP
101 Federal Street
Boston, MA 02110

William A. Streff Jr., P.C.
Michelle A. H. Francis
Craig D. Leavell
Ann H. Chen
Colleen M. Garlington
KIRKLAND & ELLIS LLP
200 East Randolph Drive
Chicago, IL 60601
(312) 861-2000 (phone)
(312) 861-2200 (fax)

By: /s/ John W. Pint
John W. Pint

EXHIBIT
15

Vehicle Navigation & Information Systems

VNIS '89

IEEE-VTS

Vehicular Technology Society

Ontario

Ministry of Transportation

Transport Canada

International Conference ✻ Toronto, Ontario, Canada ✻ September 12-14, 1989

Announcement & Call for papers — Papers are invited on vehicle navigation and information systems technology and applications, with emphasis on the topics outlined below. Presentations are welcome on user requirements, choice of technology, operations, economic assessment and performance evaluation. Of particular interest are papers on research and development programs and pilot projects, as well as those dealing with critical issues affecting system implementation such as standards, cost, market size, privacy, safety, human factors, private and public sector roles. A major goal of the Conference is to encourage interaction between the developers and potential users of this technology.

Technology

Systems

- Autonomous navigation systems
- Terrestrial and space-based radio location and navigation systems
- In-vehicle route guidance systems
- Automatic vehicle identification, monitoring and control systems
- Digital maps and geographic information systems
- Mobile data communications

User Interfaces

- Visual displays, aural communications, system controls
- Human factors considerations

Systems Analysis & Evaluation

- Simulation and graphics — traffic flow, routing, vehicle movements
- Route optimization — algorithms, static/dynamic/interactive
- Performance analysis — location, navigation, communications, control
- Lab and field test results — systems/subsystems

Technical Exhibits

A technical exhibition is planned in association with the conference. Potential exhibitors should contact the Conference organizers at the address below.

Address correspondence to:

VNIS '89 Conference
c/o Insight Planners Inc.
133 Richmond St. W., Suite 502
Toronto, Ontario
Canada M5H 2L3

Telephone: (416) 868-6565 Fax: (416) 868-0936

Applications

Driver Information

- Road and traffic conditions — pretrip and en route
- Route guidance — turn advisory, optimal routing, driver response
- Auxiliary — special data bases (roadside services, electronic "yellow pages", business, entertainment), mobile office, etc.

Fleet Management

- Monitoring — vehicle, cargo, emergency conditions, tolls, etc.
- Dispatching and routing — emergency, public and commercial fleets, dangerous goods, etc.
- Enforcement — vehicle weight, cargo, route, etc.
- Record keeping — time, distance, routes, speed, fuel.

Traffic Management

- Supply management — traffic flow and route optimization, collective/individual control
- Demand management — access control, priority treatment, restrictive zoning, road pricing, automatic billing, etc.
- Parking management — space availability and location, billing, etc.
- Data collection — traffic flow, O-Ds, trip times, incidents, etc.

Schedule

- Abstracts (500 words/6 copies) due January 16, 1989
- Notification of acceptance sent by March 27, 1989
- Camera-ready manuscript of paper due June 26, 1989
- Conference: Toronto, September 12-14, 1989.

Abstracts should clearly summarize the content of the proposed paper and contain the author's name, title, correct mailing address, telephone number and fax number. All papers will be reviewed with respect to completeness, clarity and technical soundness. Those meeting generally accepted criteria for technical papers will be published in the Conference Record.

EXHIBIT

16

FILED UNDER SEAL

EXHIBIT
17

Direction Assistance

James Raymond Davis
and
Thomas Frank Trobaugh

December 1987

Speech Research Group Technical Memo 1
The Media Laboratory
Massachusetts Institute of Technology

Abstract

Direction Assistance is an interactive program that provides spoken directions for automobile travel within the Boston area. The program has a telephone interface which uses touch tone keypad input and synthetic speech output. Routes are both short and easily followed. The directions are given in fluent English. The program has successfully directed newcomers through Boston.

This paper tells how we built Direction Assistance, with emphasis on how the available databases are and are not useful for this application. It also talks about automatic generation of route descriptions, and compares our work to that of others.

1 Introduction

1.1 Overview

Direction Assistance consists of about 11,000 lines of CommonLisp code, runs on a Symbolics Lisp Machine, and uses a Digital Equipment Corporation DecTalk synthesizer. It was written mostly during the summer of 1985 at the Thinking Machines Corporation of Cambridge, Mass. Since then, it has undergone periodic rewrites. It is running at the Media Lab, and is also installed at the Computer Museum in Boston and as part of the Age of Intelligent Machines exhibit traveling across the United States.

Direction Assistance consists of five modules. The **Location Finder** queries the user to obtain the origin and destination of the route. A location may be specified as a street address or as a telephone number. The **Route Finder** finds a simple, short route between the two points. The **Describer** generates high quality English text describing the route. The **Narrator** recites the route to the user. In addition, there is a **graphical interface** for maintenance.

These modules share a set of databases. The most important is the street map, which covers an eleven square mile area of Boston centered on the Charles River. A second database is an inverted phone directory, which provides a street address for a phone number.

In this paper, we discuss the databases, the Route Finder, and the Describer. The Location Finder and Narrator are described in [2].

It would be inappropriate to continue without mentioning the pioneering work of Jane Elliot and Mike Lesk[5,4]. Our work differs from theirs in several ways. Our interface uses synthetic speech and pushbutton telephones rather than a graphics terminal. We are much more concerned with generating fluent English text than they. On the other hand, we are not much concerned with route finding algorithms. Finally, Elliot and Lesk used a Yellow Pages database in addition to the white pages and street map. We will not clutter this paper with citations to Elliot and Lesk on every point where they have made contributions. They are to be assumed.

STREETER 00523

The New Yorker, April 6 1987 p 56

We next discuss the underlying databases, and then the modules which use them. The description of the databases will by necessity refer to features of the program in order to motivate the construction of the database.

2 Databases

2.1 Streets

Our street map began as a DIME (Dual Independent Map Encoding) file distributed by the United States Geological Survey[1]. A DIME file consists of a set of straight line segments, each with a name, a type, endpoints in longitude and latitude, and some additional information. Segment types include natural features (chiefly water boundaries), railroads, town and property lines as well as streets. The latter are also labeled with address numbers on both sides of the street at each endpoint; thus is is possible to estimate the coordinates for any street address by interpolation, assuming all lot sizes to be constant.

We began with an 11 square mile subset centered roughly on the Charles River. This includes portions of Boston (Charlestown, Allston, Back Bay, South End, North End), Brookline, and Cambridge (Cambridgeport and Harvard, Inman, Central and Kendall Squares). (See figure 1.) There are about 279 miles of streets in the map, which contains 6163 segments, of which 5506 correspond to streets. The total size is about 477 kilobytes.

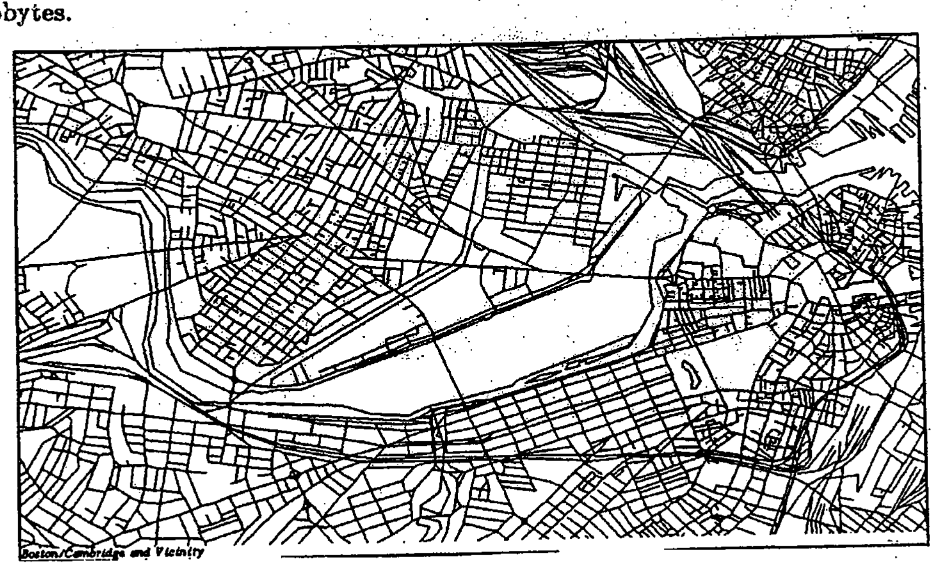


Figure 1: Street Database

The DIME file as supplied was far from suitable for our use. It contained many errors: streets were missing, mislabeled, or misconnected, and names were not spelled consistently. In some cases, more than one segment occupied the same place, and some segments were connected to themselves. We wrote a

STREETER 00525

battery of plausibility checkers to detect and remove these errors, automatically where possible.

In addition to correcting errors, we had to add new kinds of information to the database. The most important information was whether a street was one way. We also classified streets by quality, and recorded textual descriptions for some turns. We'll now describe each of these.

Segments in the DIME file are deemed to connect if they share a common endpoint. We refer to this kind of connection as ***physical connectivity***. Every segment has two endpoints, and for each of these there is a list of the segments which are physically connected to that endpoint. Obviously, physical connectivity is a symmetric non-reflexive relation. Physical connectivity is not sufficient for route finding, since it may not be legal to drive from one piece of pavement to another, even though they meet, because one might be one-way, or a turn might be forbidden, or there might be a divider in the way[1]. To provide for the fact that one can not always drive from a segment to any other physically connected to it, we added a second kind of connection, ***legal connectivity***. Two (street) segments are legally connected if one may drive from one segment to the other without breaking a law. Legal connectivity supplements, but does not replace, physical connectivity. Physically connected segments include those that can be seen in passing, and must also be retained, for they are important in forming descriptions. One cannot turn onto a railway, though the street and railroad segments are physically connected, but one may also wish to mention the crossing of the railroad as a salient detail of the tour.

Not all streets are created equal. We wanted our routes to use the widest, fastest, and most easily located streets, so we gave each street a value for goodness (super, good, average, or bad). By definition, most streets are average. The super streets are the expressways, interstate highways, and other limited access roads. Our rating of super is awarded more on the basis of being easy to find and to follow, since super roads are often crowded and slow. At the other extreme, the bad streets are those we know to be narrow or in poor repair. Our database contains only three miles of such streets. Unlike the taxi driver, we are not interested in shortcuts which use marginal streets.

The concept of "better than average" is a bit hard to define. We wanted to identify streets which were likely to be easy to find and follow. We decided that streets that were long were likely to be important, so we marked all streets longer than one half mile as "good", and then added a few more by hand if they seemed important. The resulting network is about 105 miles long, and forms a simplified skeleton covering our map. It appears in figure 2.

The third extension was to expand the street classification scheme. We added new segment types for bridges, underpasses, rotaries, and access ramps. This information is useful to both the route finder and to the describer, as we show below.

Finally, at every intersection in the map we can store additional descriptive information about each possible turn at the intersection, in the form of labelled items. Each item has a label telling what kind of information is stored, for instance an exit number or the text of a sign present at that intersection. This information is used by the Describer.

We made almost all of these corrections and augmentations ourselves from observations in the field.

[1]In this case, the turn is forbidden by physical obstacles, and not merely law or custom. But rather than engage in an epistemology of barriers, we use the same mechanism to represent this restriction.

Figure 2: Network of good streets

We could not find a paper map listing all the one way and restricted turning streets of Boston, so we had to drive around looking for them. This investment in time and effort is a major cost of the system, but needs to be done only once. The graphic database editor was extremely useful, as it permitted rapid editing of the database. We commend the many designers of the Lisp Machine window system for making this easy.

2.2 Neighborhoods

A related database lists the neighborhoods of Boston, with their associated zip codes. We need this database because a given street might occur in several different towns. For instance, there are three distinct streets named "Washington" in our map, in Boston, Cambridge, and Somerville. Even worse, Cambridge contains two different streets named "Elm".

The Location Finder uses this database to disambiguate street names. When the user supplies a name that could designate more than one street, it is necessary to ask for further information, e.g. "Do you mean Beacon Street in Cambridge or in Boston?". To make this as easy as possible, it is best to use the names of the most general locations that still distinguish the streets[2]. If the street occurs in two neighborhoods of the same city, the neighborhood name is used. If the street occurs in different cities, the city name is sufficient. We determine neighborhood from the Zip code of the street. The mapping from Zip to neighborhood is imperfect, but good enough for our purposes. For the most part, the neighborhood names are those used by the local post offices. We think it is very likely that these names are also familiar to the local residents, and intelligible to visitors, but we have no evidence.

[2]This assumption could be tested. If people represent locales hierarchically, and if there is a preferred level of representation, it might be *more* difficult to determine inclusion in a too-general region.

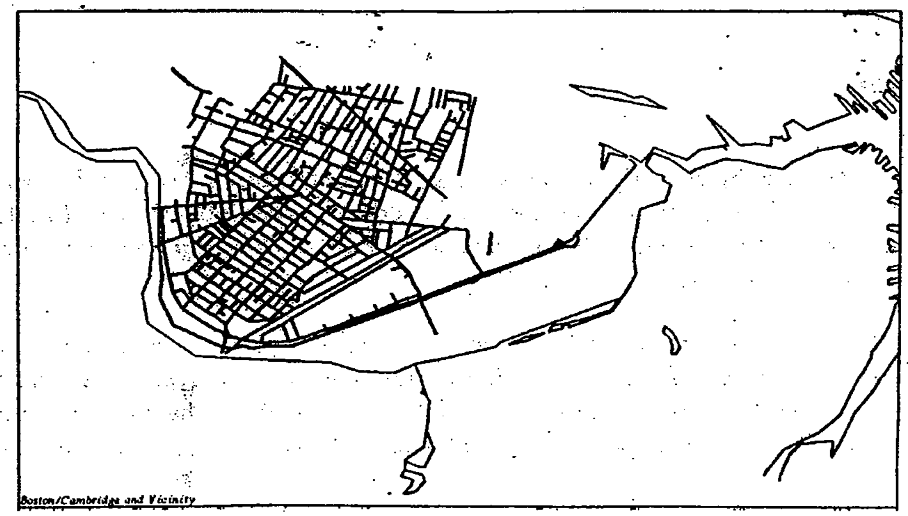


Figure 3: Central Square, 02139

2.3 Inverted Phonebook

The inverted telephone directory allows us to map telephone numbers to street addresses. We built this database ourselves, by inverting a "white pages" database. This required parsing the street addresses in the white pages. which was difficult for several reasons. The white pages have a great variety of spelling and abbreviation. We found, for instance, 23 variations of "Massachusetts". In addition, the format is not consistent. Sometimes listings contain professions ("atty" or "archt"), or a second phone number ("If No Answer"), or other information (e.g. "toll free", "children's phone"). We did not have the typographic information that helps separate names from locations and phone numbers. Finally, addresses are often incomplete, listing only a city, or road, or some a name which does not correspond to a street, such as a shopping center or an office park.

Even after parsing, it can be hard to determine locations from a a phone book listing. Even the best entries have at best a street, number, and city. But as we said above, streets occur in more than one place within a city. There is a rough correspondence between exchange and locale, so we can sometimes determine a unique location with this extra information. But when we can not, the Location Finder must ask the user to choose a location, as it does for street names.

Having described the databases, we now turn to the modules of Direction Assistance.

3 Route Finder

The Route Finder finds a route subject to three constraints. The route must be easy to follow, reasonably short, and it must be found before the user loses patience[3]. These constraints conflict. Rarely is there a straight line route - the shortest route may require devious shortcuts. We are biased towards simplicity, since we want our users not to get lost.

The output of the Route Finder is a *path*, an ordered list of street segments, such that the origin is on the first segment, the destination on the last, and each segment is legally connected to the next. The real time requirements of the system rule out exhaustive, breadth first search[4]. The current implementation uses a best first search that provides reasonably good routes in a moderate time. A sample route appears in figure 4.

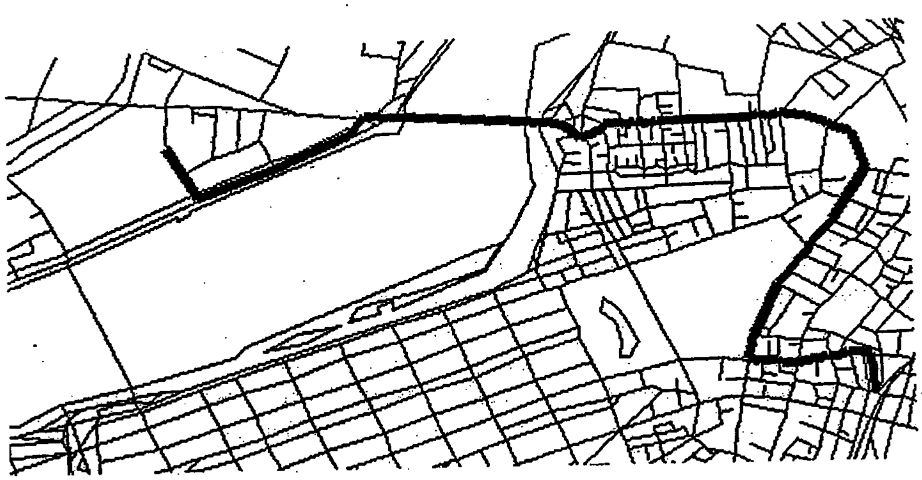


Figure 4: A sample route

Best first search is an improvement on breadth first search. Search is conducted in (simulated) parallel on a list of candidate partial paths. For each path, there is a cost which is the sum of the known cost for the current path and an estimate which is a lower bound on the cost for (as yet undetermined) remainder of the path. At each step of the search, we consider the path of least cost, and expand it by considering all segments legally connected to its terminal end. The estimation function is just the Cartesian distance, since no route can be shorter than a straight line. Figure 5 shows every segment visited by the search in finding the route shown above.

As Elliot and Lesk point out, it is not desirable to find minimum distance routes, for these have too

[3]A fourth constraint which we do not consider explicitly is that the route must be easy to describe. We are familiar with situations where a person asks for a route to a familiar place, but we can not describe the route because it is a "felt path": we no longer remember (or do not know) the names of the streets, only a list of subtle cues we can't describe.

[4]on a serial machine, anyway. An experimental version on the Connection Machine[6] works in just this way.

STREETER 00529

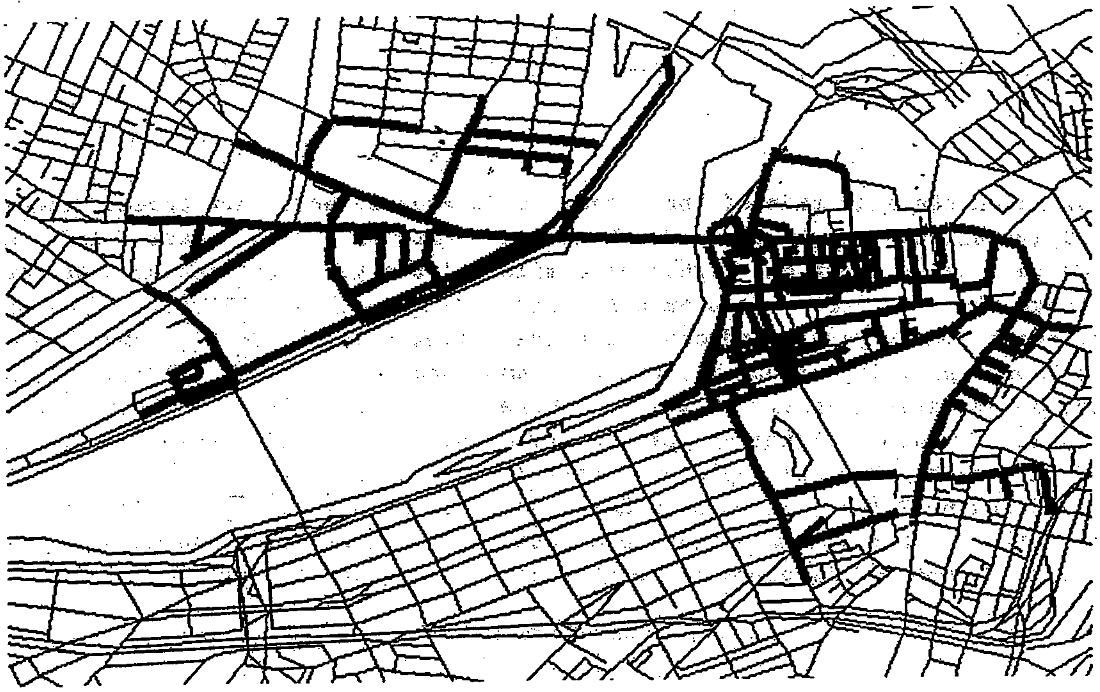


Figure 5: All segments touched by search

many turns. Such routes are hard to describe and hard to follow. Elliot and Lesk impose a cost of 1/8 mile for a right turn, and 1/4 mile for a left turn. We extend their system of costs in several ways. First, we consider street goodness. Travel down a "super" street is not as "expensive" as travel down an average street, and travel down a "bad" street incurs a surcharge. Second, we consider sharp right turns to be as bad as left turns, since they are harder to execute. Third, we reduce or waive turn costs in some cases. For example the turning cost is halved for a turn on to or off a one-way street, and waived altogether for a forced turn ("left turn only"). A turn onto a bridge is also free, since bridges are major landmarks, and contribute to ease of following the route. We have not studied the effect of these routes on the routes found, nor have we attempted to determine whether the routes are better where different. Such a study would require a model of driver's errors, both of understanding and of execution.

4 Describer

The Describer generates a set of text instructions for following the route. (An example of its output appears in figure 6.) We generate text instead of a map for two reasons. First, the system is used by telephone, which limits the output to voice. But even if our users had portable graphics terminals with modems, we would prefer text to graphics, because some people can not read maps. In a survey of map reading abilities Streeter and Vitello recommend text as a "lowest common denominator"[9].

The Describer creates a new representation of the route, instead of using the path itself. There are two reasons for this second form of representation. First, the elements of a a path (segments) are too fine grained for useful textual description. Recall that a segment reaches from just one intersection to the next. This is smaller than our sense of a "street", which continues as a unity past many intersections. In addition, segments are straight lines: so a street with no intersections might be still represented as a

```
If your car is on the same side of the street as 20 Ames Street, turn around, and start
driving. Drive all the way to the end, about one eighth of a mile. Make a left onto
Memorial Drive. Drive about one eighth of a mile. After you pass Wadsworth Street on
the left, take the next left. It's an easy left. Merge with Main Street. Stay on Main
Street for about ninety yards, and cross the Longfellow Bridge. You'll come to a
rotary. Go about half way around it, and turn onto Cambridge Street. Drive all the way
to the end, about three quarters of a mile. Make a right onto Tremont Street. Drive
about one half of a mile. After you pass Avery Street on the left, take the next left
onto Boylston Street. Stay on Boylston Street for about one eighth of a mile. After
you cross Washington Street, it becomes Essex Street. Keep going. Drive about one
eighth of a mile. After you pass Ping On Street on the right, take the next right onto Ed-
inboro Street. Number 33 is about one eighth of a mile down on your right side.
```

Figure 6: sample of directions

sequence of segments if it made a broad turn. We want to describe the entire stretch of a street as a single object. A second reason is that a path is just a topological structure, but natural instructions should be expressed in terms of geometry and of types of streets. Consider the difference between a "fork", a "T", and an "exit", as shown in figure 7. All have the same topology - a branch in the road. But they must be described differently. The Describer's structure is a *tour*, which is a sequence of *acts* to be taken in following the path.

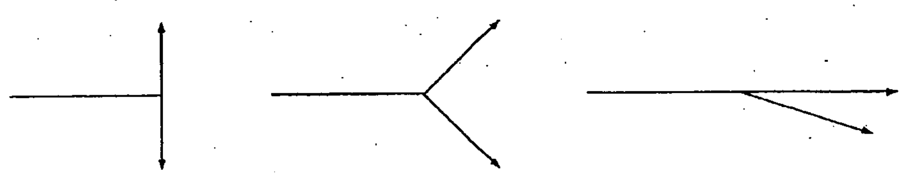


Figure 7: T, fork, and exit all have same topology

4.1 Acts

Acts are things a driver does (or notices) while following a route. Figure 8 shows our taxonomy of acts.

Each of these acts must be *recognized*. The route finder works only with segments, and the Describer builds acts which describe motion from segment to segment. We now describe each of these acts, and how we recognize them. We describe the text generated for each below.

STREETER 00531

- Boundaries
    - Start
    - Stop
- Straight
    - Name Change
- Turn
    - Enter
    - Exit
    - Merge
    - Fork
    - U Turn
    - Rotary
    - Ordinary

Figure 8: Act Taxonomy

The first act is necessarily START, and the last STOP. They are trivial to recognize. The NAME CHANGE act requires the driver to notice a change in name, but nothing further. We include it only to avoid confusion. The difference betweeb a NAME CHANGE and a TURN is that the former consists of a two streets meet within 10 degrees of straight, and where there is no other segment at the intersection with the same name as either of them. These two criterion are almost correct, but not quite right. There are streets which seem (to us) to be name changes, but have more extreme turns (at least, as represented in the map). For the present, we have caused these to be treated as name changes by changing the map, slightly altering the positions to make the turns more gentle. This would be intolerable were we using the map for, say, surveying, but is of no consequence for route description.

There are several types of TURN acts. The ENTER and EXIT acts refer to limited access roads. In this case, some of the travel will often be on "nameless" segments - access ramps. This shows one reason for the additional classification of street segments. We want to recognize entrances and exits, and we want to describe access ramps in different terms than other streets.

A MERGE and a FORK are similar in that they are different actions that might be taken at the same intersection, depending upon the direction one is driving. A Merge has the following characteristics:

1. Old and new streets have different names.
2. Only one street is legally possible.

3. The angle of turning is small.
4. There is at least one other street going to the destination street.
5. All streets make only small turns onto the destination.

At a FORK on the other hand, there are at least two ways to go, though all are shallow turns. Note that a "fork" onto an exit ramp is recognized as an EXIT.

There are two types of U turn known to drivers in Boston. The first kind is made in the middle of the street (within a single segment, in our representation). Our routes never include such turns. Not only are they illegal, such moves never shorten the path. The second kind of U turn is the sort one makes to reverse direction on a divided road. Typically one makes a left onto a nameless piece of road, which is often very short, and then makes a second left. This double turn is what we call a U TURN act. It is very important to recognize this act, because describing it as two successive lefts is very confusing. It is a single entity in the minds of drivers. We recognize a U TURN as a pair of turns where the intermediate segment is less than 165 yards long, the total angle is within 20 degrees of 180, and the name of the street is unchanged after the two turns.

Perhaps the most insidious feature of Boston's streets is the ROTARY. For those not familiar with the term, a rotary is a one way street in a circle. Traffic enters the rotary on roads which are (usually) tangent to the circle, moves counterclockwise around the circumference, and exits on another tangent. Rotaries are difficult to traverse because they cars enter and exit within a very short distance, without much room to maneuver. Recognition of a rotary is trivial, but only because we label all rotary segments explicitly in the street map.

An ORDINARY turn is anything not handled by one of the above cases.

4.2 Cues

While the Describer is collecting the acts of the tour, it also collects **cues**. A cue helps the driver follow the tour. We distinguish four kinds of cues. *Action* cues tell when to do an act. *Confirmatory* cues describe things that will be seen while following the route. *Warning* cues caution the driver about possible mistakes. A warning successfully heeded also serves as a confirmatory cue. *Failure* cues describe the consequences of missing an act, e.g. "If you see this, you have gone too far".

The most common action cue is just the name of the street. An instruction such as "Turn right onto Tremont Street." tells the driver what to do and when to do it. This cue may be hard to follow, since street signs may be missing. A very strong action cue is coming to the end of a road. No one is likely to forget to turn under this circumstance, since the alternative is to leave the road. We refer to this as a "forced turn" cue.

Distance traveled is also a cue, but hard to use. People have a vague sense of distance, but not an accurate one. Still, we use distance as a secondary cue, because we can compute it easily and it helps some people. We express distance in yards when less than 1/16 of a mile, and other distances in approximate

fractions of a mile because people are accustomed to seeing distances expressed this way. We do not use tenths of miles, because some people do not know how to use odometers, and because using an odometer to calculate distance requires doing mental arithmetic, which might prove distracting while driving.

We never use blocks, since a block is not a clearly defined concept. We do not know whether a block is bounded by an intersecting street, or only by streets that cross and continue. Figure 9 illustrates this. In

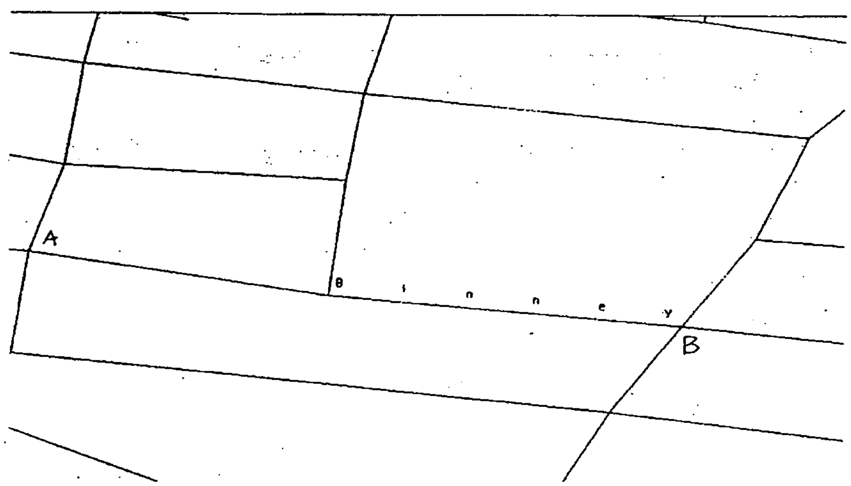


Figure 9: Is the distance between "A" and "B" one block, or two?

any event, we do not expect our drivers to be able to drive more than two or three blocks without losing count. Since we don't want to rely on distance or counting blocks, we use as a cue for an act the name of a street immediately preceding the act. This is a risky cue, since the driver who misses the cue may keep looking for it and miss the destination street as well. To make this less likely, we use only streets on the same side as the turn for a cue. This way, a driver need attend to only one side of the road while looking for street signs, so if the cue street is missed, the target street may still be seen. This same strategy is adopted in [10].

The confirmatory cues are crossing major streets or railroads, or going through an underpass. The only warning cue currently is a warning about left exits from limited access roads. We assume drivers will not take the wrong exit, but if they keep in the left lane they may be surprised by an unexpected left exit. We have not implemented failure cues.

4.3 Generating Text

For each act there is a corresponding routine which generates one to three sentences describing it. The routine selects appropriate cues from all those gathered. Now we'll describe some aspects of generating text.

```
(defun disc-seg-rotary (act)
  (list
   (make-sentence
     "You'll" "come" "to"
     (make-np-constituent '("rotary") :article :indefinite))
   (make-conjunction-sentence
     (make-sentence
       "Go" (rotary-angle-amount (get-info act 'rotary-angle))
       "way" "around" (make-anaphora nil "it"))
     (make-sentence
       "turn" "onto" (make-street-constituent (move-to-segment act) act)))))

(defun rotary-angle (angle)
  (selector angle <=
    (45 '("just" "a" "little"))
    (135 '("about" "a" "quarter"))
    (225 '("about" "half"))
    (315 '("about" "three" "quarters" "of" "the"))
    (360 '("almost" "all" "the"))))
```

Figure 10: Generator for rotary

Generating text for a START is tricky because it is hard to specify an initial direction. We do not use absolute directions, because most people do not know them. If we had a landmark database we might sometimes use relative direction (e.g. "towards the river"). Instead, we use the initial address, since that also determines a side of the street, and thus a direction to drive. We might have used "If your car is on the same side of the street as ... start driving the way it is facing.", but that sounded clumsy. Instead, we chose to give a negative instruction, either "If your car is on the same/opposite side of the street from ... turn around, and start driving." For one way streets we mention that the street is one way, and say "just start driving." We think (but do not know) that drivers would not have confidence in the instruction ("just start driving.") if it did not indicate that the system knew about the one way street.

One of the simplest generators is for rotaries. It appears in figure 10. Rotaries are hard to describe and hard to follow, because there are no good references for distance around a rotary. We can not expect people to measure angular distance around the rotary, and there may not be signs. The segments of a rotaries may or may not be nameless, or there may be several names involved. The rotary itself may have a name, e.g. Leverett Circle, but this name does not appear in the database and usually does not appear on any street signs either.

Output from this generator appears in figure 6. The generator produces two sentences, the second of which is a conjunction of two sentences. The distance around the rotary is converterted from an absolute angle, as measured on the map, to an approximation in English.

The instructions generated have syntactic structure only for sake of exploiting generality in text generation. Thus the function `make-np-constituent` handles agreement between the article and the noun. The function `make-sentence` ensures that capitalization and punctuation are correct. Text is sent directly to the synthesizer, and punctuation is required to achieve proper intonation. The function `make-anaphora` serves no purpose at present, but in planned future research will allow us to convey intonational features of discourse[3].

4.4 Comparison

We can compare our descriptions with those generated by Streeter and colleagues[10].

Streeter's descriptions are intended to be understood and acted upon in real time, as if uttered by a navigator in the next seat. (In fact, they are recorded on a tape, and the driver pushes a button to play the next instruction.) This interface imposes a new requirement on the form of the directions. Since they are to be heard and acted on in real time, it is important to repeat essential information so that it can be remembered. In our interface, we assume people are writing down the directions before they begin to drive, so repetition is not crucial. (The user can ask the Narrator to replay an instruction if it is not understood.)

They classify turns into ordinary turns, "T" turns, complex intersections, turns in short succession, and continues. Their "T" turn is our "forced turn" cue. The difference between an ordinary turn and a "T" turn is that the latter needs no failure cue. So our treatments are similar. We do not distinguish complex intersections, though we should. The Route Finder should avoid them, and the Describer should warn about them.

Their instructions are sometimes more structured than ours. They cluster turns which occur close to each other into a single instruction block, and their "continue" is just our "name change", but is also incorporated into the following turn. We recognize the importance of providing higher levels of structure, and wish to remind the reader that Streeter and company were working by hand, not with a program, and were in a better position to form hierarchies than we.

We claim that our directions are more natural than those of Elliot and Lesk, but have no proof for this. We leave it to the reader to judge.

Are our directions clear? We know that people have been able to follow them, but we have not made any systematic test. Christoper Riesbeck wrote a program (MCMAP) which judged the clarity of directions. Our directions would not be acceptable to it, to judge from its published description. Partly this is because we talk about features the program does not know, for example rotaries, but also because his program explicitly rejects use of miles for distance as inherently unclear. We use mileage only as an approximation, as a cue for when to look for a landmark, but the weak syntactic powers of MCMAP would not notice this. Also, we use "drive all the way to end", which Riesbeck terms a "procedural operator", and did not implement. Since people accept our directions, this suggests that Riesbeck's rules are too strict, or perhaps not powerful enough.

5 Discussion

Products like Direction Assistance are beginning to appear in the marketplace. It is reported that ETAK, of Sunnyvale California, has a product (the Navigator) which, installed in a car, estimates the car's position by counting wheel rotation and turning angle and comparing results with a stored map. A display in the dashboard displays the local area and the position of the car. The Navigator does not supply driving directions, but surely could be made to do so.

A more similar product is DriverGuide, made by Karlin and Collins, also of Sunnyvale, which is reported to produce printed directions for travel in the Bay Area[8].

5.1 Better databases are required

Any serious use of Direction Assistance requires further improvements to the street map. The area covered is too small, and even the small region covered is not fully mapped. More significantly, there are additional facts that the current street database format can not represent.

Among these are time-dependent legal restrictions (e.g. "no left turn during rush hour"), restriction of height, weight, and prohibition of commercial vehicles, multiple names of streets, presence of stop lights, and landmarks. In addition, the representation of addresses is not sufficient. We have seen addresses with fractions and with letters, and there are also streets where both even and odd numbers are on the same side of the street.

A practical system must account for multiple names. When Route 93 passes through Boston, it is also Route 3, the Fitzgerald Expressway, and the Southeast Expressway. When Massachusetts Avenue turns north at Harvard Square, only the southbound lane is "really" Massachusetts Avenue. The other direction is officially Peabody Street. We do not know which name to use when naming these streets, but we should at least be able to accept all synonyms on input.

Boston, like any city, changes its configuration of streets daily. Some changes, e.g. for construction, are temporary, although they may persist for years. Others are permanent. Streets are built and removed, and sometimes they change names or directions. A practical system requires accurate and timely corrections to the database.

We could give better directions with a better database, giving, for example the location of traffic lights or landmarks such as gas stations. Elliot and Lesk were able to capture business locations from an online Yellow Pages. To be more ambitious, we might hope for a representation rich enough to capture the qualities of image and orientation described by Kevin Lynch[7]. We have no proposal for how to do this at present.

5.2 Applications

We initially designed Direction Assistance with tourists in mind. Boston's confusing streets often lead the visitor astray. A tourist's direction guide could be provided by the city, or as a profitable venture. But a tourist may not know the street address or phone number for the destination. In fact, there may not be one, for the destination might be a general area, such as a neighborhood or park. Tourists would probably prefer to identify locations by name. It might be difficult to add this feature without making the interface more complicated.

Direction Assistance could direct people to services. Given the caller's location and the type of service desired, Direction Assistance could select the closest, and provide a route. This service might be dedicated to a single vendor (e.g. for banking machines) or as an advertising service for many customers.

Routing delivery vehicles pose special problems. Some of the most useful routes in Boston are closed to commercial vehicles, either for legal reasons or because they have such low underpasses that even scofflaws can not get through. We could extend the street database to record such facts.

We also feel compelled to mention the implications of Direction Assistance for privacy. Should a public Direction Assistance include home telephone numbers? People may want to keep the ability to give out their home phone numbers without also revealing their addresses to callers. One can hang up on an annoying caller. A visitor may be harder to dispose of.

Acknowledgments

A prototype version was written by Dinarte R. Morais during the winter of 1985. We are indebted to him for decoding the DIME files, the initial window system interface, and the proof of concept. We made extensive use of a database package and string matcher written by Craig Stanfill. Charles Lieserson made major improvements to the search algorithm of the Route Finder. Fletch McCellan of the PhoneBook Corporation loaned us the raw phone book database. This work would not have happened without the guidance and persistence of Brewster Kahle. This paper was much clarified by the comments of Janet Cahn, Mike Hawley, Margaret Minksy, and Chris Schmandt. We thank them all.

This work was supported at MIT by the DARPA Space and Naval Warfare Systems Command, under contract numbers N00039-89-C-0406 and N00039-86-PRDX002 and by the Nippon Telegraph and Telephone Public Corporation. Hardware support was provided by Symbolics and Digital Equipment Corporation.

This paper bears the names of two authors, for the program was joint work. But though it is written in the plural, it is the work of only one of us. I dedicate it to Tom, who did not survive to see his work described. Though too small a memorial, it is the best I can manage at this time.

References

[1] *Geographic Base File GBDF/DIME: 1980 Technical Documentation.* U.S. Department of Commerce, Data Users Services Division, 1980.

[2] James R. Davis. Giving directions: a voice interface to an urban navigation program. In *Proceedings of 1986 Conference*, pages 77–84, American Voice I/O Society, Sept 1986.

[3] James R. Davis and Julia Hirschberg. Automatic generation of prosodic support for discourse structure. In *Proceedings of the Association for Computational Linguistics*, page (submitted), 1988.

[4] R. J. Elliot and M. E. Lesk. *Let Your Fingers Do the Driving: Maps, Yellow Pages, and Shortest Path Algorithms.* Technical Report unpublished, Bell Laboratories, 1982.

[5] R. J. Elliot and M. E. Lesk. Route finding in street maps by computers and people. In *Proceedings of the National Conference on Artificial Intelligence*, pages 258–261, 1982.

[6] W. Daniel Hillis. *The Connection Machine.* MIT Press, 1985.

[7] Kevin Lynch. *The Image of the City.* MIT Press, 1960.

[8] Ronald Rosenberg. Mapping out a new idea. *The Boston Globe*, 39, 1987. February 17.

[9] Lynn A. Streeter and Diane Vitello. A profile of drivers' map reading abilities. *Human Factors*, 28:223–239, 1986.

[10] Lynn A. Streeter, Diane Vitello, and Susan A. Wonsiewicz. How to tell people where to go: comparing navigational aids. *International Journal of Man/Machine Systems*, 22(5):549–562, May 1985.

STREETER 00539

EXHIBIT
18

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF MASSACHUSETTS

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, **Plaintiff,** **v.** **HARMAN INTERNATIONAL INDUSTRIES, INCORPORATED,** **Defendant.**	**Case No: 05-10990 DPW** **Hon. Douglas P. Woodlock**

MIT'S FIRST SUPPLEMENTAL RESPONSE
TO HARMAN'S INTERROGATORY NOS. 11-15

Pursuant to Rules 26 and 33 of the Federal Rules of Civil Procedure, Plaintiff, Massachusetts Institute of Technology ("MIT") submits the following supplemental responses and objections to Harman International Industries, Incorporated's ("Harman's") Interrogatory Nos. 11-15 (the "Interrogatories").

GENERAL OBJECTIONS

The following general statements and objections are incorporated into each of MIT's responses, as set forth there in full, even if not repeated therein:

1. MIT objects to Harman's method of counting, but under Harman's counting method, Harman has served sixty-one (61) interrogatories, and thus has exceeded the twenty-five (25) permitted by Rule 33 of the Federal Rules of Civil Procedure. Upon mutual agreement of counsel as to method of counting the parties' respective interrogatories, MIT will provide additional appropriate responses, if necessary.

2. MIT objects to the Interrogatories to the extent they call for disclosure of information protected by the attorney-client privilege, work-product doctrine, and/or any other privilege or immunity. In the event that any response given by MIT contains privileged or protected information, its disclosure is inadvertent and shall not constitute a waiver of any privilege or protection with respect to the divulged information or any other information.

3. MIT objects to the Interrogatories to the extent they attempt or purport to impose obligations on MIT beyond those required by the Federal Rules of Civil Procedure 26 and 33, or the Local Rules of Practice of the United States District Court for the District of Massachusetts.

4. MIT objects to the Interrogatories to the extent that they seek information already in Harman's possession, equally available to Harman, and/or publicly available.

5. MIT objects to the Interrogatories to the extent that they seek information not in the possession, custody, or control of MIT, information not owned or belonging to MIT, or information that is subject to a non-disclosure obligation pursuant to a confidentiality agreement with a third-party.

6. MIT objects to the Interrogatories to the extent that they are vague, ambiguous, and/or confusing, incomprehensible and/or unanswerable because of undefined or ill-defined terms and/or confusing syntax, or they fail to describe with reasonable particularity the information sought.

7. MIT objects to the Interrogatories to the extent that they are overly broad, unduly burdensome, oppressive, and/or designed solely to harass MIT.

8. MIT objects to the Interrogatories to the extent that they seek information not relevant to the subject matter of the present lawsuit and/or are not reasonably calculated to lead to the discovery of admissible evidence.

9. The information supplied in MIT's responses may not be based solely upon the knowledge of the executing parties, but may include the knowledge of MIT's agents, representatives, and attorney(s), unless privileged.

10. MIT expressly reserves all objections as to relevance and/or admissibility of any information disclosed in its objections and/or responses.

11. MIT's willingness to provide responses to any of the Interrogatories is not a concession that the subject matter of the particular Interrogatory is discoverable, relevant to this action, or admissible as evidence.

12. To the extent MIT adopts any terms used by Harman in its Interrogatories, such adoption is specifically limited to the objection and responses herein, and does not constitute an admission of law or fact by MIT.

13. The presence or absence of any general or specific objection does not mean that MIT does not object on any other grounds.

14. MIT has responded to the Interrogatories as it interprets and understands each Interrogatory made therein. If Harman subsequently asserts an interpretation of any Interrogatory that differs from the understanding of MIT, MIT reserves the right to supplement its objections and responses.

15. MIT incorporates its General Objections to Harman's First and Second Sets of Requests for the Production of Documents and Things (Nos. 1-29; 30-61) as if fully set forth herein.

16. The responses set forth below are based on information presently known to MIT. MIT expressly reserves the right to complete its investigation and discovery of the facts and to rely, at the time of trial or in other proceedings, upon documents and evidence in addition to the information provided regardless of whether such information is newly discovered or currently in

existence. MIT may, in the future, obtain or locate additional information responsive to these Interrogatories. Further, a complete response to certain Interrogatories depends in part upon information to be adduced from Harman or third parties during discovery. MIT, therefore, reserves its right, at any time, to revise, amend, correct, supplement, modify, or clarify its responses, on a timely basis, in accordance with Federal Rules of Civil Procedure 26 and 33.

SPECIFIC OBJECTIONS AND RESPONSES

INTERROGATORY NO. 11

The Schmandt and Davis, "Synthetic Speech for Real Time Direction-Giving" publication (MIT 01101-02) notes "field trials" of the Back Seat Driver that occurred more than 1 year before the filing date of U.S. Patent No. 5,177,685. Mr. Davis' thesis (*see* HAR 001479) also notes that the Back Seat Driver had been used more than 1 year before the filing date of U.S. Patent No. 5,177,685. For each asserted claim of United States Patent No. 5,177,685, identify each and every limitation of the claim that MIT contends was not embodied in a field trial prior to August 9, 1989, and explain in detail all bases for any contention by MIT that such field trials do not render each asserted claim of the '685 patent invalid under 35 U.S.C. § 102(b).

SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 11

MIT objects to this Interrogatory as overly broad, unduly burdensome, and not reasonably calculated to lead to the discovery of admissible evidence. MIT further objects to this Interrogatory because it calls for a legal conclusion with respect to validity. MIT further objects to this Interrogatory because it seeks information protected by the attorney-client privilege, work product doctrine, and/or other applicable privileges or immunities.

Subject to and without waiving the foregoing general and specific objections, MIT states that the field trials do not constitute invalidating prior art under 35 U.S.C. § 102; but to the extent

Harman contends they do relate to prior art, MIT will timely submit the report(s) of its expert(s) with respect to validity, and thus intends to supplement this Interrogatory in a timely manner, after completion of any such report(s). MIT further incorporates by reference the deposition testimony of Dr. James R. Davis, Ph.D. and Christopher M. Schmandt in response to this Interrogatory, where questions related to this line of interrogatory were answered.

MIT further states that any field trials that occurred prior to the filing date of the '685 patent do not constitute a public use, anticipatory or otherwise, under U.S. patent law. To the extent that Harman contends that field trials do constitute public use or sale of the Back Seat Driver system, MIT further responds that the field trials were solely for experimental purposes, performed under both the control and supervision of Dr. Davis and Mr. Schmandt, and did not disclose to the public the claimed invention. Any field trials involved Dr. Davis or Mr. Schmandt observing operation of and driver response to the Back Seat Driver. The results observed by Dr. Davis and Mr. Schmandt were not disseminated except under the control of Dr. Davis and Mr. Schmandt to a limited group of people. In addition to MIT's general incorporation by reference above, MIT identifies Page 29, Line 3-Page 33, Line 4; Page 83, Line 17-Page 99, Line 11 of Dr. Davis' deposition testimony.

INTERROGATORY NO. 12

To the extent MIT contends there exist any secondary considerations or indicia that support a finding of non-obviousness of any claim(s) of U.S. Patent No. 5,177,685, identify all alleged bases for any such contention and all documents or other materials that support MIT's contention.

SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 12

MIT objects to this Interrogatory as overly broad, unduly burdensome, and not reasonably calculated to lead to the discovery of admissible evidence. MIT further objects to this Interrogatory because it calls for a legal conclusion with respect to obviousness, non-obviousness and/or secondary considerations. MIT further objects to this Interrogatory because it seeks information protected by the attorney-client privilege, work product doctrine, and/or other applicable privileges or immunities.

Subject to and without waiving the foregoing general and specific objections, MIT states that it will timely submit the report(s) of its expert(s) with respect to validity, and thus intends to supplement this Interrogatory in a timely manner, after completion of any such report(s). At this time, it appears that there are numerous objective indicia of non-obviousness, discussed in detail in the several depositions taken to date, including long felt, but un-met need as expressed by, for example, various prophetic articles cited by Harman, which describe the wishes of many for a working product, but which singularly or in combination fail to disclose or suggest a working embodiment of the claimed invention as well as commercial success by the accused infringer.

Further objective indicia of nonobviousness include the commercial success in both the United States, Europe, Japan, and elsewhere, by Harman and by others, of navigation systems that are covered by the claims of the '685 patent, as evidenced by the substantial sales and the market share of such systems.

The failure of other alleged prior art systems to gain market share further illustrates the nonobvious nature of the '685 patent, especially, as Harman and others, currently sell systems that are covered by the '685 patent, rather than systems described in the prior art. Moreover, the failure of digital map-makers such as Navigation Technologies to combine digital map

functionality with discourse into an operational system despite industry-leading capabilities in digital map-making underscores the leap achieved by the '685 patent.

The '685 patent was based on pioneering work worthy of a doctoral thesis from one of the most prestigious engineering schools in the world, and is the subject of tribute from skilled scientists in the fields of, for example, computational linguistics and automobile navigation systems. For example, Barbara Grosz, professor at Harvard and pioneer in the field of computational linguistics teaches graduate-level courses based on Dr. Davis' thesis and the '685 patent.

Moreover, within the portfolio of patents owned by the MIT Media Lab, MIT understands that the '685 patent is the patent most often cited, both by Patent Examiners and by patent applicants, in general, and specifically relating to computational linguistics and automobile navigation.

INTERROGATORY NO. 13

For each claim of U.S. Patent No. 5,177,685, identify the date(s) on which the subject matter recited therein was first completely conceived, and identify by Bates number all documents or other material that evidence all such date(s) in any way.

SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 13

MIT objects to this Interrogatory as overly broad, unduly burdensome, and not reasonably calculated to lead to the discovery of admissible evidence. MIT further objects to this Interrogatory as premature to the extent that it calls for a legal conclusion with respect to conception. MIT further objects to this Interrogatory to the extent that it mischaracterizes the legal standard for conception. MIT further objects to this Interrogatory because it seeks

information protected by the attorney-client privilege, work product doctrine, and/or other applicable privileges or immunities.

Subject to and without waiving the foregoing general and specific objections, MIT objects to the phrase "completely conceived", but states that the subject matter of the '685 patent was conceived before the filing date of the application on which the '685 patent issued. The details of the conception were fully described in answer to numerous questions to the inventors propounded during the deposition testimony of Dr. James R. Davis, Ph.D. and Christopher M. Schmandt and in response to this Interrogatory, and those answers are herein incorporated by reference.

MIT further states that claims 1, 3, 4, 7-10, 14, 16, 19, 24, 27, 28, 42-44, 46, 48, 55, and 57-58 were conceived at least as early as April 1988. MIT further states that claims 2, 5, 6, 11-13, 15, 17, 18, 20-23, 25, 26, 29-41, 45, 47, 49-54, and 56 were conceived at least as early as June of 1989.

INTERROGATORY NO. 14

For each claim of U.S. Patent No. 5,177,685, identify the earliest date(s), if any, on which the subject matter recited therein was first actually reduced to practice, and identify by Bates number all documents or other material that supported all such date(s).

SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 14

MIT objects to this Interrogatory as premature to the extent that it calls for a legal conclusion with respect to reduction to practice. MIT further objects to this Interrogatory to the extent that it mischaracterizes the legal standard for reduction to practice. MIT further objects to this Interrogatory because it seeks information protected by the attorney-client privilege, work product doctrine, and/or other applicable privileges or immunities.

Subject to and without waiving the foregoing general and specific objections, MIT states that the subject matter of the '685 patent was reduced to practice before the filing date of the application on which the '685 patent issued. The details of the reduction to practice were fully described in answer to numerous questions to the inventors propounded during the deposition testimony of Dr. James R. Davis, Ph.D. and Christopher M. Schmandt and in response to this Interrogatory, those answers are herein incorporated by reference.

MIT further states that claims 1-12, 15, 19, 20, 23-28, 30-49, 51, 52, and 54-58 were reduced to practice at least as early as June 1989. MIT further states that claims 13, 14, 16-18, 21, 22, 29, 50, and 53 were reduced to practice no later than the filing date of the '685 patent.

INTERROGATORY NO. 15

For each claim of U.S. Patent No. 5,177,685, explain in detail (including an identification by Bates number of all documents that evidence in any way) all alleged diligence by or on behalf of Mr. Davis and/or Mr. Schmandt in reducing to practice the subject matter recited in the claim.

SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 15

MIT objects to this Interrogatory as overly broad, unduly burdensome, and not reasonably calculated to lead to the discovery of admissible evidence. MIT further objects to this Interrogatory as premature to the extent that it calls for a legal conclusion with respect to diligence and/or reduction to practice. MIT further objects to this Interrogatory to the extent that it mischaracterizes the legal standard for diligence and/or reduction to practice. MIT further objects to this Interrogatory because it seeks information protected by the attorney-client privilege, work product doctrine, and/or other applicable privileges or immunities.

Subject to and without waiving the foregoing general and specific objections, the details of the diligence towards reduction to practice were fully described in answer to numerous

questions to the inventors propounded during the deposition testimony of Dr. James R. Davis, Ph.D. and Christopher M. Schmandt and in response to this Interrogatory, those answers are herein incorporated by reference.

See, e.g., MIT 00933-42, MIT 01374-78, and throughout MIT 00457-MIT 00621.

Dated: April 27, 2006

Respectfully submitted,

Massachusetts Institute of Technology,

By its Attorneys,

Kimberly A. Mottley

Steven M. Bauer (BBO# 542531)
Kimberly A. Mottley (BBO# 651190)
John W. Pint (BBO# 660548)
PROSKAUER ROSE LLP
One International Place
Boston, Massachusetts 02110-2600
Phone: 617-526-9600
Fax: 617-526-9899

CERTIFICATION

I, the undersigned, have reviewed MIT's First Supplemental Responses to Harman's Interrogatory Nos. 11-15. The responses set forth herein, subject to inadvertent or undiscovered errors or omissions, are based on and therefore necessarily limited by the records and information still in existence, presently recollected, thus far discovered in the course of preparation of the responses, and currently available to MIT. Consequently, MIT reserves the right to make any changes in or additions to any of these responses if it appears at any time that errors or omissions have been made therein or that more accurate or complete information has become available. Subject to the limitations set forth herein, said responses are true to the best of my present knowledge, information and belief.

I hereby certify under penalty of perjury that the foregoing is true and correct.

Executed on this __th day of April, 2006.

____

John H. Turner, Jr.
Associate Director, Technology Licensing Office
On behalf of Massachusetts Institute of Technology

CERTIFICATE OF SERVICE

I HEREBY CERTIFY that on April 27, 2006, I caused a true and correct copy of MIT's FIRST SUPPLEMENTAL RESPONSES TO HARMAN'S INTERROGATORY NOS. 11-15 (NOS. 8-20) to be served on the following counsel of record via email:

Robert J. Muldoon, Jr.
SHERIN AND LODGEN, LLP
101 Federal Street
Boston, MA 02110

William A. Streff Jr., P.C.
Michelle A. H. Francis
Craig D. Leavell
Ann H. Chen
Colleen M. Garlington
KIRKLAND & ELLIS LLP
200 East Randolph Drive
Chicago, IL 60601
(312) 861-2000 (phone)
(312) 861-2200 (fax)

Kimberly A. Mottley

Kimberly A. Mottley

EXHIBIT 19

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF MASSACHUSETTS

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, **Plaintiff,** **v.** **HARMAN INTERNATIONAL INDUSTRIES, INCORPORATED,** **Defendant.**	**Case No: 05-10990 DPW** **Hon. Douglas P. Woodlock**

MIT'S SECOND SUPPLEMENTAL RESPONSE
TO HARMAN'S INTERROGATORY NOS. 13-14

Pursuant to Rules 26 and 33 of the Federal Rules of Civil Procedure, Plaintiff, Massachusetts Institute of Technology ("MIT") submits the following supplemental responses and objections to Harman International Industries, Incorporated's ("Harman's") Interrogatory Nos. 13-14 (the "Interrogatories").

GENERAL OBJECTIONS

MIT herein incorporates by reference its General Objections as set forth in MIT's First Supplemental Response to Harman's Interrogatory Nos. 11-15.

SPECIFIC OBJECTIONS AND RESPONSES

INTERROGATORY NO. 13

For each claim of U.S. Patent No. 5,177,685, identify the date(s) on which the subject matter recited therein was first completely conceived, and identify by Bates number all documents or other material that evidence all such date(s) in any way.

SECOND SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 13

MIT objects to this Interrogatory as overly broad, unduly burdensome, and not reasonably calculated to lead to the discovery of admissible evidence. MIT further objects to this Interrogatory as premature to the extent that it calls for a legal conclusion with respect to conception. MIT further objects to this Interrogatory to the extent that it mischaracterizes the legal standard for conception. MIT further objects to this Interrogatory because it seeks information protected by the attorney-client privilege, work product doctrine, and/or other applicable privileges or immunities.

Subject to and without waiving the foregoing general and specific objections, MIT objects to the phrase "completely conceived", but states that the subject matter of the '685 patent was conceived before the filing date of the application on which the '685 patent issued. The details of the conception were fully described in answer to numerous questions to the inventors propounded during the deposition testimony of Dr. James R. Davis, Ph.D. and Christopher M. Schmandt and in response to this Interrogatory, and those answers are herein incorporated by reference.

MIT further states that claims 1, 3, 4, 7-10, 14, 16, 19, 24, 27, 28, 42-44, 48, 55, and 57-58 were conceived at least as early as April 1988. MIT further states that claims 2, 5, 6, 11-13, 15, 17, 18, 20-23, 25, 26, 29-41, 45-47, 49-54, and 56 were conceived at least as early as June of 1989.

MIT identifies documents bearing Bates numbers MIT00433-MIT00947, MIT01101-MIT01102, MIT01370-MIT01378, MIT01955-MIT02002, and MIT02155-MIT02274 as responsive to this Interrogatory. MIT further identifies pages 83-90, 104, 145-162, 179, and 287 of Mr. Schmandt's deposition transcript as responsive to this Interrogatory. MIT further

identifies pages 79-91, 129, 167-169, 205, and 219-238 of Dr. Davis' deposition transcript as responsive to this Interrogatory.

INTERROGATORY NO. 14

For each claim of U.S. Patent No. 5,177,685, identify the earliest date(s), if any, on which the subject matter recited therein was first actually reduced to practice, and identify by Bates number all documents or other material that supported all such date(s).

SECOND SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 14

MIT objects to this Interrogatory as premature to the extent that it calls for a legal conclusion with respect to reduction to practice. MIT further objects to this Interrogatory to the extent that it mischaracterizes the legal standard for reduction to practice. MIT further objects to this Interrogatory because it seeks information protected by the attorney-client privilege, work product doctrine, and/or other applicable privileges or immunities.

Subject to and without waiving the foregoing general and specific objections, MIT states that the subject matter of the '685 patent was reduced to practice before the filing date of the application on which the '685 patent issued. The details of the reduction to practice were fully described in answer to numerous questions to the inventors propounded during the deposition testimony of Dr. James R. Davis, Ph.D. and Christopher M. Schmandt and in response to this Interrogatory, those answers are herein incorporated by reference.

MIT further states that claims 1-12, 15, 19, 20, 23-28, 32-49, 51, 52, and 54-58 were reduced to practice at least as early as June 1989. MIT further states that claims 13, 14, 16-18, 21, 22, 29-31, 50, and 53 were reduced to practice no later than the filing date of the '685 patent.

MIT identifies documents bearing Bates numbers MIT00433-MIT00947, MIT01101-MIT01102, MIT01370-MIT01378, MIT01955-MIT02002, and MIT02155-MIT02274 as

responsive to this Interrogatory. MIT further identifies pages 83-90, 104, 145-162, 179, and 287 of Mr. Schmandt's deposition transcript as responsive to this Interrogatory. MIT further identifies pages 79-91, 129, 167-169, 205, and 219-238 of Dr. Davis' deposition transcript as responsive to this Interrogatory.

Dated: May 2, 2006

Respectfully submitted,

Massachusetts Institute of Technology,

By its Attorneys,

Steven M. Bauer (BBO# 542531)
Kimberly A. Mottley (BBO# 651190)
John W. Pint (BBO# 660548)
PROSKAUER ROSE LLP
One International Place
Boston, Massachusetts 02110-2600
Phone: 617-526-9600
Fax: 617-526-9899

CERTIFICATION

I, the undersigned, have reviewed MIT's Second Supplemental Responses to Harman's Interrogatory Nos. 13-14. The responses set forth herein, subject to inadvertent or undiscovered errors or omissions, are based on and therefore necessarily limited by the records and information still in existence, presently recollected, thus far discovered in the course of preparation of the responses, and currently available to MIT. Consequently, MIT reserves the right to make any changes in or additions to any of these responses if it appears at any time that errors or omissions have been made therein or that more accurate or complete information has become available. Subject to the limitations set forth herein, said responses are true to the best of my present knowledge, information and belief.

I hereby certify under penalty of perjury that the foregoing is true and correct.

Executed on this __th day of May, 2006.

____

John H. Turner, Jr.
Associate Director, Technology Licensing Office
On behalf of Massachusetts Institute of Technology

CERTIFICATE OF SERVICE

I HEREBY CERTIFY that on May 2, 2006, I caused a true and correct copy of MIT's SECOND SUPPLEMENTAL RESPONSES TO HARMAN'S INTERROGATORY NOS. 13-14 to be served on the following counsel of record via email:

Robert J. Muldoon, Jr.
SHERIN AND LODGEN, LLP
101 Federal Street
Boston, MA 02110

William A. Streff Jr., P.C.
Michelle A. H. Francis
Craig D. Leavell
Ann H. Chen
Colleen M. Garlington
KIRKLAND & ELLIS LLP
200 East Randolph Drive
Chicago, IL 60601
(312) 861-2000 (phone)
(312) 861-2200 (fax)

John W. Pint

EXHIBIT

20

Jim Davis
jrd3@alum.mit.edu
Toronto, Canada

Professional

I'm interested in building software systems that improve communication among people. I believe that communication mediums of the future will have an increasing understanding of the structure and content of the messages they transmit. They will manipulate, reformat, and even generate that content. I am interested in hypertext systems, network information access, and collaboration.

Lately, I'm interested in Ruby.

Technical Biography

The first image I had of computers came from my father, when he was a field repairman for IBM. He used to bring home broken boards for me to play with. The PC (printed circuit) boards of those long ago days were a lot more interesting to look at than today's boards, they had lots of different shapes and colors. Of course I had no clue what any of them did. In 1973 I entered MIT, where I studied at the Architecture Machine Group, the ancestor of the Media Lab. From 1977 until 1985 I worked in a variety of industrial and research positions, including the Multics project, and culminating in the Atari Cambridge Research Lab. In 1985 I entered the doctoral program at the Media Lab. My main pieces of work there were *Direction Assistance* (which gives spoken driving instructions over the telephone) and *Back Seat Driver* (which does the same thing in a car, while you're driving.). In 1989 I received a PhD, and after a year's post-doc I began to work for Xerox, first at the Design Research Institute, and then at PARC. Since then, I have worked for two startups (Coursenet and Intelligent Markets) in SF, with a brief stop at Sybase in between. At this moment, I'm working as a software consultant.

What's here

collected papers
: The papers I've published in the technical/scholarly world.

unpublished writings

software
: Free, unsupported sofware, all of which has been released by Xerox, but is provided with NO WARRANTIES, EXPRESS OR IMPLIED, of any kind. Includes, among others, a cheap HTML parser in perl.

Hacks, including SuDoKu solver in Javascript and Conway's game of Life in Javascript.

Pictures, including Death Valley (2002), Yosemite (2003), and Patagonia (2006).

Personal Biography

This is brief, because all I really hope to do here is distinguish myself from all the other people named "Jim Davis". There are a lot of us. In particular, I am not the one who draws the comic strip featuring the cat (whose name I will not mention here, lest it make my web page even more likely to be found by those seeking that person.)

I was born in Berkeley, California on December 28, 1955. I went to high school in Suffern, New York (40 miles north of the city). In 1973 I moved to Cambridge to attend MIT. My son Adam was born in 1977, the same year I graduated. In 1985 I returned to MIT to get my PhD. In 1990 I married Anna Korteweg. In 1991 I moved to Ithaca, New York, and in 1996 back to Berkeley, where my daughter Michal Korteweg Davis was born on May 28, 2002. In July, 2004, we moved to Toronto, Canada, where I am now a landed immigrant. My second son, Ruben Daniel Korteweg Davis, was born on October 10, 2006.

I played in some rock bands in Boston and Ithaca, New York. All were fun, none were famous.

I used to practice the dance form known as Contact Improvisation, and for many years was the curator of the CI web site, but no longer. It is now at http://www.contactimprov.net/.

Okay, am I the Jim Davis you were looking for? If not, good luck searching for the right one.

If you want my technical, academic, or industrial history, then please read my resume.

Updated 28 Feb 2007

EXHIBIT

21

Abstract

The Back Seat Driver is an automobile navigation aid which uses synthetic speech to give driving instructions in real time to the driver of a car. The advantage of speech over visual aids is that it leaves the driver's eyes free for driving, however it also poses special problems. This paper describes the strategies employed by the Back Seat Driver to successfully use speech. We hope this paper will persuade you of the value of speech in driving directions.

Introduction

The Back Seat Driver uses synthetic speech to give driving instructions in real time to the driver of a car. Speech is the only output channel it uses. There are no graphics. This paper discusses the advantages and problems arising from our exclusive use of speech to provide directions. The first section presents a brief overview of the Back Seat Driver. The second section describes the linguistic abilities of the Back Seat Driver. The final section describes the problems we have encountered because of our exclusive use of speech, and how we have overcome them.

System Overview

The architecture of the Back Seat Driver is shown in figure 1. At the center of the Back

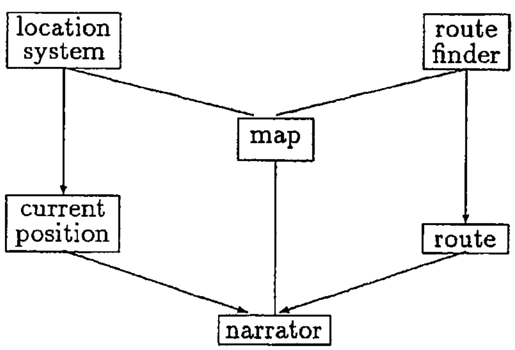


Figure 1: Back Seat Driver components

Seat Driver is the map database. The street map represents two ways in which streets can

be connected: *physical* connectivity means it is physically possible to drive from one segment to another, and *legal* connectivity means it is lawful to do so. Legal connectivity is obviously needed to find legal routes, and physical connectivity for correctly describing intersections. The street map also includes traffic lights, stop signs, the number of lanes, and the location of all gas stations. These features are useful for both route finding (since, e.g. fast routes should avoid traffic lights) and for descriptions. The location system (supplied by the project sponsor, NEC) determines the current position of the vehicle by dead reckoning and map matching. It is further described in [?]. The driver gives the Back Seat Driver a destination by entering an address on a keyboard. Using this map, the route finder can find the shortest route, the simplest one, or the one most easily followed, depending on the driver's preference.

The *narrator* is the subject of this paper. It generates instructions spoken by a speech synthesizer (a Dectalk). The narrator follows the driver's progress along the route. It decides what to say by comparing the current position against the map. The system follows the driver's progress, giving each instruction just when needed. If the time between instructions is long, the program gives the instruction twice, first in a detail, and later in a brief form. When not otherwise occupied, the system may deliver voice mail messages, weather reports, or commentary about the route. If the driver makes a mistake the system automatically finds an alternate route and continues.

The system has been running in prototype form since April 1989. It has been successfully used by drivers who have never driven in Boston. A somewhat longer description of the system appears in [?]. A complete description of the system appears in [1].

Linguistic Abilities

In designing the Back Seat Driver we chose to use speech as the sole means of providing driving instructions for two reasons. First, we believe that the driver's eyes are already employed watching traffic, and best left undisturbed. Second, we know that the alternative (map displays) will not work for those people who have difficulty reading maps[3]. We were also influenced by an experiment on route following which compared spoken instructions with paper maps[4]. Subjects who heard spoken directions did better than those with maps, and also better than those with *both* sources of guidance. Although this experiment does not compare real time speech to real time maps, it does suggest that spoken directions might be easier to follow than visual directions.

Classifying Actions

Based on a study of how people naturally give spoken driving instructions, we developed a taxonomy of intersection types (Figure 2). This taxonomy is necessary in order to describe an intersection in the same way that a person would. For example, people talk about a "T"

turn differently than a "fork" (or "Y") in the road. It is important that instructions match people's perceptions of the the world they see.

- CONTINUE
- FORCED-TURN
- TURN-AROUND
- TURN
- FORK
- ENTER
- EXIT
- ONTO-ROTARY
- EXIT-ROTARY
- STOP

Figure 2: Act taxonomy

The proper classification of an intersection depends upon the topology (how many streets are at an intersection), the geometry (the angles among them), and the types of roads involved. For instance, the difference between the "T" and "fork" mentioned above is one of geometry, not topology (figure 3), and the difference between a "fork" and an exit from a highway is that one of the two roads in the "Y" of the exit is much larger than the other.

Figure 3: A "T" and a "Y" have the same topology

In our system, a route is a sequence of street segments leading from the origin to the destination. We consider every connection from one segment to another as an "intersection",

even if there is only one next segment at the intersection. At any moment, the car will be on one of the segments of the route, approaching an intersection (unless an error occurs, which is handled as discussed below). The task of the Back Seat Driver is to say whatever is necessary to get the driver to go from the current segment, across the intersection, to the next segment of the route.

The items in the taxonomy of intersection types are called **acts**. We use an object oriented programming methodology, so for each act there is a corresponding "expert". The Back Seat Driver generates speech by consulting these experts. At any moment, there will be exactly one expert in charge of telling the driver what to do. To select this expert, the Back Seat Driver asks each expert in turn to decide whether it applies to the intersection. The experts are consulted in a fixed order, the most specific ones first. The first expert to claim responsibility is selected. This expert then has the responsibility of deciding what (if anything) to say.

Describing actions

Each expert is able to generate text which describes the intersection. A description for an act must tell the driver two things: what to do and when (or where) to do it. "What to do" is expressed by a more or less constant verb phrase which depends upon the taxonomic classification, but may also depend upon specifics of the intersection. Thus a slight turn might be described by the verb "bear" where a sharper turn would be a "turn". The descriptions can be verbose or brief, and they can be expressed in past, present, or future tense. (We'll say why this flexibility is needed below.)

Saying "when"

Our study of natural instructions showed us that people almost never use distance as a cue for when to act. This is in sharp contrast to the textual directions provided by systems such as that of the Hertz rental company. Instead, people use two strategies. They wait until the driver is close to the intersection before saying anything, and/or they use a great variety of landmarks – including traffic lights, stop signs, other signs, buildings, road features, and the positions of other moving objects (e.g. "Follow that car."). The Back Seat Driver adopts both of these strategies.

Speech is especially useful as a cue for timing because speech is a temporal event, with a clear beginning and ending time. You know when someone begins to speak and when they finish. Someone peering at a map displayed on a CRT may have trouble distinguishing two adjacent streets, but there is no mistaking the word "now". Using time as a cue minimizes the workload on the driver, because the navigator absorbs the burden of remembering when to act. It also demands that the navigator have an accurate idea of where the car is. Our

system demands positional accuracy of no greater than 10 meters for successful operation.

The Back Seat Driver's use of landmarks is unique in vehicle navigation systems. Our database began as a DIME file, but we extended it to include traffic lights, stop signs, road features (such as overpasses, bridges, and tunnels), distinctive signs, and the location of gas stations. Most of these are represented as attributes of the segments in the map database. To select a landmark for an intersection, the Back Seat Driver looks backwards from the intersection for the closest landmark which is also unique – that is, it prefers to say "take the first right after the underpass" rather than "take a right at the second set of lights". We think this makes the landmark easier to remember.

The Back Seat Driver does not speak at every single intersection. In the great majority of cases, it is perfectly obvious to the driver what to do (namely, to continue on forward). The action experts are also capable of deciding when the action at the intersection should be obvious to the driver. At present, the only action that is *ever* treated as obvious is CONTINUE. It is usually obvious to continue across an intersection, but we have found that what is obvious to one driver may not be so to another. Some people, for instance, are not comfortable driving across a major intersection unless they are instructed to do so. The expert can be somewhat customized so that its judgment of "obviousness" will correspond to that of the driver. If the action at the next intersection is obvious, the Back Seat Driver says nothing about it, and looks ahead for action at the next intersection, until it finds one that is *not* obvious.

The Back Seat Driver gives instructions just prior to the action. It also gives instructions further in advance, if time permits. This is especially useful when the instructions are complicated, as they are at some intersections. It is also able to give instructions "on demand". We call this the "what now" button. Drivers use this button for two reasons. Sometimes they are unsure whether they have come to the place where they are supposed to act, so they press the button to find out. At other times, they reach an intersection where the Back Seat Driver says nothing, because it believes the action is obvious, but it is not obvious to the driver. When the driver hits the "what now" button, the expert for the upcoming intersection describes it, even if it is considered to be obvious.

Talking about past and future

An advantage of language over pictures or gestures is that it can express events in the past or future. This advantage is well appreciated by readers of fiction, but may not yet be appreciated by designers of navigation systems. A navigation system should be able to talk about the past and future of the route, not just the present.

Drivers often need advance notice to prepare for an action. An example is what we call **lane advice**, which tells the driver to get into, or stay out of, a given lane. Lane advice is common in natural directions, and is one of the most appreciated features of the Back Seat Driver.

One reason for talking about the past is to describe mistakes. Drivers do not always follow the route the Back Seat Driver intends, either because of a mistake by the driver, the program, or external circumstances. When a mistake occurs, the Back Seat Driver finds a new route from the current location to the destination, while the driver is still moving. It also describes the mistake, saying something like "Oops, I meant for you to go straight." We think it is important that the system tell the user that there has been a mistake (without casting any blame on the user!) so that the user will come to better understand the system's style of instruction giving, and so that the user will remain confident in the system's understanding of the route. Talking about past and future actions is important in navigation. Speech seems to be the easiest way of doing this.

Example

As an example, here's a sample of the description of the left turn from Fulkerson Street to Main Street in Kendall Square, Cambridge.

> Get in the left lane because you're going to take a left at the next set of lights. It's a complicated intersection because there are two streets on the left. You want the sharper of the two. It's also the better of them. After the turn, get into the right lane.

This description was generated by the TURN expert in verbose form. It begins with some lane advice, then specifies the next action and provides a landmark for the place. The turn is described, and the proper street is described by two independent cues, one geometric, and one qualitative. Finally, the text provides a second piece of advice for after the turn.

Summary

The speech interface of the Back Seat Driver provides instructions without requiring the driver to look away from the road. Using speech permits us to talk about the past and the future as well as the present, and to give more detailed descriptions of the act than are possible with maps. Furthermore, it allows us to specify timing with great precision. But speech is not without its problems. The next section will discuss them, and the steps we have made to overcome them.

Liabilities of Speech

The advantages of a spoken language interface, as described above, do not come without cost. First, there are problems common to any natural language interface: while it is not

EXHIBIT
21

(Part 2)

terribly difficult to make a rudimentary interface, language generation requires substantial programming effort to be fluent and natural. Language is complicated, and people have literally a lifetime of experience with it, and are sensitive to fine nuances. On the other hand, having made this effort, we can exploit these nuances to to convey extra information.

A second problem is that a natural language interface is only useful to those who speak the language. In our experience, only a few non-native speakers have been able to understand the directions. Map displays have conventions of their own, but are more universal than natural language. We have also noticed that some driving terms used in the Boston area (e.g. "rotary") are not in the dialect of other English speakers. In our view, universality is not a prime concern. We believe that systems should be custom fit to the idiosyncrasies of their owners. The Back Seat Driver in your car should speak to you in the language and terms that are best for you as an individual, not you as a generic human. Equality is for politics, not interfaces.

The remainder of this section discusses problems specific to *spoken* natural language generation.

Speech takes time

As we said above, speech is inherently temporal. We take advantage of this when we use speech as a timing cue, but it also can be a difficulty. A real time spoken navigation system must plan its speech to ensure that it has enough time to say what it needs to say. If little time remains, it must say less (or speak more quickly), or ask the driver to slow down. We handle this problem by tracking the vehicle's position and velocity, and by modeling the time required to speak. The Back Seat Driver begins its speech at a time chosen to be early enough to allow the driver to hear the entire message, understand it, and react to it, before the point where action must be taken. The model of reaction time includes a constant for the driver's comprehension and a variable time which depends on the speed of the car, according to the maximum comfortable braking deceleration.

The temporal nature of speech also requires that the Back Seat Driver sometime combine instructions into a single utterance. When uttering an instruction, the Back Seat Driver looks ahead for the next instruction. If it determines that the time between the end of the execution of the current instruction and the beginning of the next is too short to allow it to speak the next instruction, it combines that text into the current one.

The Back Seat Driver does more than just give directions. Among other things, it also reads electronic mail messages from our office, gives weather reports, and makes comments about the route and road. Because speech takes time, and because a spoken utterance is only useful if completely spoken, the Back Seat Driver must carefully allocate the right to speak among potential tasks. It is undesirable for one task's speech to interrupt another's.

Speech can be misunderstood

A liability of speech, and synthetic speech in particular, is that speech can be misunderstood. This is particularly a problem with street names, because there are constraints that can help a driver correct a partially misunderstood name. A driver hearing an utterance that sounds like "Tarn reft" can guess that it is a corrupt form of "Turn left", but nothing can help the driver know what was intended by "Tarn Street". Directions should not use street names, because street name signs may be hard to see, misaligned, or simply missing. The importance of this first became apparent when we observed one driver who consistently misunderstood names, but also did not realize that he had misunderstood. Furthermore, the strength of his faith in the name was so strong that he drove straight through intersections, despite being told to "take the next left". This is probably the right thing to do with human instructions, where names are usually correctly understood, but street counts (e.g. "the third right") are imprecise or simply wrong. Our directions are phrased to minimize the use of street names in instructions. A typical text is: "Take the second left. It's Franklin Street."

Speech is transient

Information presented by speech does not persist, except in short term memory. We have already mentioned this as a reason why instructions should be given as late as possible. Another consequence of the transience of speech is that the system must be able to repeat itself at any time, since the driver may not always be able to hear the speech. Repetition is not literal, in contrast to, say, a program which reads the newspaper aloud, because the situation changes with time. For instance, if asked to repeat "Take the third left", the system may instead say "Take the second left" if the car has crossed an intersection. The consequence for the implementation is that the system retains not its previous words, but rather the previous reason for speaking. When asked to repeat, it invokes the same function that produced the last utterance. Depending on the circumstance, this function may repeat the same words, or may choose a different wording.

A second problem with the ephemeral quality of speech is that the driver has no evidence of the program's existence except when it is speaking. We consider it very important that the driver have continued confidence that the program is running correctly, is aware of the driver's position and progress, and is "seeing" the world in the same way the driver does. We have devoted substantial effort to maintaining the illusion of **co-presence**.

In the introduction to this section, we said that the nuances of language could be used to convey much information. Co-presence is an idea communicated more by nuance than by explicit statement. (People would laugh if the system said "I'm right here with you." It sounds like something a therapist would say.) One way we indicate co-presence through

nuance is by using **deictic** pronouns. Deictics are words that "point" at something. In English, we have four deictic pronouns: "this", "that", "these", and "those". The first two are singular, the second plural. The difference between "this" and "that" (and "these" and "those") is that "this" refers to something close. We use this in referring to landmarks. When the landmark is close, we use the proximal form (e.g. "these lights") ; when distant, we use a brief noun phrase (e.g. "the next set of lights"). This is important. When a driver is stopped 30 meters back from a stop light, it may be literally true to say "turn left at the next set of lights", but it will confuse the driver.

A second means of conveying co-presence is to acknowledge the driver's actions. After the driver carries out an instruction the system briefly acknowledges the act if there is time, and if the act was not so simple (e.g. continuing straight) as to need no acknowledgment. This acknowledgment is a short phrase like "Okay". Some drivers dislike acknowledgments, so they can be disabled, but most find the confirmation comforting. The timing of the acknowledgment does much to confirm the driver's sense that the program really knows where the car is. Another source of acknowledgment is the use of **cue words** in the instructions. It will often by the case that the route calls for the driver to do the same thing twice (e.g. make two left turns). The speech synthesizer we use has very consistent pronunciation, and drivers sometimes get the impression that the system is repeating itself because it is in error (like a record skipping). The acknowledgments help to dispel this, but we also cause the text to include cue words such as "another". These indicate that the system is aware of its earlier speech and the driver's previous actions.

Yet another means of conveying co-presence is to make occasional remarks about the road and the route. These remarks indicate that the program is correctly oriented. As an example, when the road makes a sweeping bend to one side, the program speaks of this as if it were an instruction ("Follow the road as it bends to the right.") even though the driver has no choice in what to do. The program also warns the driver about potentially hazardous situations, such the road changing from one-way to two-way, or a decrease in the number of lanes. As with acknowledgments, these warnings can be disabled if the driver dislikes them. Other remarks have less to do with the route. We justify these by the maxims of cooperative conversations formulated by philosopher H. P. Grice[2]. His maxim of QUANTITY (part 1) says: "Make your contribution as informative as is required." Grice explains that one can convey information by appearing to flout the maxim. In this case, a driver can reason as follow: "The program, like all cooperative agencies, obeys the maxim of quantity. Therefore, it is had something important to say, it would say it. The program said nothing of great significance, therefore there is nothing urgently requiring my attention. So everything is well." At present, our "Gricean" utterances are trivial observations about the weather, but we are re-designing them to convey useful information about the city.

Summary

A speech interface for giving driving instructions has several advantages over a graphics interface. There are problems with natural language interfaces in general, and speech in particular, but they can all be overcome. The result is an excellent aid for navigation.

Acknowledgments

The authors wish to gratefully acknowledge the support of NEC Home Electronics, Ltd.

References

[1] James R. Davis. *Back Seat Driver: voice assisted automobile navigation.* PhD thesis, Massachusetts Institute of Technology, September 1989.

[2] H. P. Grice. Logic and conversation. In Cole and Morgan, editors, *Syntax and Semantics: Speech Acts*, volume 3, pages 41–58. Academic Press, 1975.

[3] Lynn A. Streeter and Diane Vitello. A profile of drivers' map reading abilities. *Human Factors*, 28:223–239, 1986.

[4] Lynn A. Streeter, Diane Vitello, and Susan A. Wonsiewicz. How to tell people where to go: comparing navigational aids. *International Journal of Man/Machine Systems*, 22(5):549–562, May 1985.

EXHIBIT

22

A voice interface to a Direction giving program

James Raymond Davis

January 1987

Speech Research Group Technical Memo 2
The Media Laboratory
Massachusetts Institute of Technology

Abstract

This paper describes the user interface to Direction Assistance, a program which provides high quality directions for driving between two points in the Boston area. The interface employs synthetic speech and touch tone input to provide access to telephone callers. The interface is able to diagnose common user errors and explain itself to the user. This paper also discusses some of the problems of synthetic speech for this type of interface. The most significant limitation of presently available speech synthesizers is the lack of ability to specify prosodic features.[*1].

Direction Assistance

Direction Assistance [3] is an program that provides directions for automobile travel within the Boston area. The program has a detailed street map covering about 11 square miles, with about 279 miles of streets, including bridges, tunnels, and one way streets. A routing algorithm produces routes that are both short and simple. A text generation program provides natural language descriptions of these routes. A graphics interface displays routes on a map. The task of the voice interface is to make this power available to naive users through the telephone, by using synthetic speech and and touch tone keypad.

The overall structure of Direction Assistance is shown in figure 1.

The *Location Finder* obtains the locations of the start and destination, the *Route Finder* produces a route which is both direct and easily followed. The *Describer* produces a tour, a description of the route, and the *Narrator* recites the tour to the user in manageable chunks. The Location Finder and Narrator are the modules that communicate with the user. Both use a common set of interface routines to provide for a uniform interface.

The Location Finder and the Narrator differ in their behavior, because the Location Finder is input oriented, while the Narrator is output oriented. The Location Finder needs to obtain two locations, and it does this by asking questions and soliciting input. The user can specify location by giving either a street number and name, or by giving a telephone number (the program's data base includes an inverted phone book). It needs to get several different types of input - answers to yes or no questions, telephone numbers, street numbers and names. It has little to say, other than to guide the user through the input process.

The Narrator, on the other hand, has a lot to say, and needs no input from the user. Its main concern is flow control, making sure that it is going as fast as the user can write, but no faster. Both of these modules also have the secondary goals of keeping the user oriented within the interaction and recovering from errors

The initial message

Direction Assistance is designed to be used by people with no experience with computers or synthetic speech. Almost every American is at least familiar with the concept that computers might talk, but they are not experienced with interfaces built with present day technology. There are lots of ways to go wrong. The interface tries to forestall some of

hem with its initial message:

```
This is Direction Assistance. I can speak to you, but I can not hear your voice. To
tell me anything, you must use the keypad on your telephone. If you get confused any
time while you're using me, hit the key with the star or asterisk,
```

The initial message conveys two important ideas. First it tells users how to communicate. Until recently, people expected to use the telephone to talk with people. Nowadays, people are beginning to get used to having "conversations" with answering machines, which speak, and then listen. Very few people know how to interact with a program that speaks often, but listens only for button pushes.

The second idea is that there is always help available, and always in the same way, by hitting the "star" key. This is a surprisingly difficult concept to convey. Examination of the log shows that many of users never used the help key, and this was even true for people who daily used computer systems that had "Help" keys. There are several reasons this might be so. First, this initial message is unexpected, and may have caught people by surprise. Second, the message has no direct relevance to the task at hand, and people may have been too impatient to listen to it. Much of the testing was done with a much longer message (nearly one minute) and this surely taxed people's concentration. Finally, people may not have understood what "star" meant. It is hard to find names for the keys without numbers. Some people prefer "asterisk". The "number sign" key is even worse, being called "pound sign", "sharp sign", "hash" or even "tic tac toe" by different people.

Different types of input

There are four types of input Direction Assistance needs from the user: answers to yes or no questions, selection from a small list of items, numbers, and names. We now consider each in turn.

Yes or No

The interface poses a yes or no question by stating a possibility and requesting the user to hit any key if it is true, thus for example "If you want to enter a different address, please hit any key now." If the user takes no action within a sage always includes the consequence of the "no" if it is not obvious. This protocol is easier to use than one that requires an explicit action for either "yes" or "no", but can be slower, since the user has to wait for the full duration allowed if the answer is no. This duration must be long enough for the user to make a decision.

Selection

Selection means choosing one of a small list of items. There are two types of selection. In a sequential selection the system slowly reads a list of choices, pausing briefly after each, until the user designates one by striking a key. This is effectively a yes or no question ("Do you want this one?") for each element of the list. In an enumerated selection the system reads the entire list, assigning each item a number, then asks the user to hit the key for the number of the desired item. Each type of selection has its advantages and disadvantages,

Sequential selection is simple to use but can be very slow, since the user may have to listen to several undesired choices before hearing the one wanted. A second problem is that the user may not know which choice is the best without hearing them all. For this reason, if the user makes no selection, the selection routine offers to repeat the list,

Enumerated selection is harder to use than sequential selection because the user must remember and enter a number. The compensation for this complexity is that it can be faster than sequential selection. There are two reasons for this. First, the list is read faster, since the system need not pause to await a user action while reading. More importantly, if the user already knows the order of the options she can enter the number immediately, without waiting for the list to be read. An enumerated selection is appropriate when the same list of choices will be presented more than once.

Numbers

MIT 02256

There are two types of numbers in this application, phone numbers and street address numbers. Phone numbers are entered in the familiar manner. The system simply collects the next seven button pushes. Street numbers require a delimiter, since the number of digits is not predictable. The delimiter is the number sign, since the only other key is the help key. Both phone numbers and street numbers require some additional checking, described below.

Names

The user selects street names by spelling them with the letters on the keypad. There are three reasons this is difficult: First, people are not familiar with the layout of the letters - they have to "hunt and peck" for letters. Second, the letters Q and Z, and the space character are missing. These are assigned to the 1 key, which bears no letters[*2]

The third problem is that each of the keys has three different letters, and there is no obvious way for the user to designate which of the three is intended. Other systems have resolved this problem by using more than one button push per character, either with a shift sequence (e.g using keys *, 0, and # to select one of the letters.) or by hitting a key N times to select the N'th letter upon it. Direction Assistance is able to bypass this problem because the set of all street names in Boston is much smaller than the set of all digit sequences. A given digit sequence will almost always specify only one street name. If it does not (e.g. "MILL" and "MILK") the system asks the user to make a sequential selection among them. Witten [13] found 8% collisions in a 24,500 word dictionary when using this approach. In this application there are much smaller number of collisions, just eight out of 1000 names. This is both because of the smaller size of the set and because street names are longer than dictionary words. The mean length of a street name is seven characters.

Absence of Context is a problem

A telephone interface is harder to use than a graphics interface (such as a form filling environment or a menu) because there is no context present while the user is entering data. The interface always prompts for input, but the user may not understand the prompt because of noise, inarticulate speech, or inattention. Even if the prompt is correctly heard, it must also be remembered, for it is not present during the time the user is entering the data. In a graphics environment the prompt remains visible.

The problems, then, are that the user may not know what to enter or how to delimit it. Neither of these are explicitly indicated, other than in the prompt.[*3] If a closing delimiter is expected, and none arrives after a fixed delay, a timeout occurs, and the interface reminds the user that a delimiter is required.

Handling Hangup

The interface is more complicated than it needs to be, because there is no way to detect whether the caller has hung up the phone. A user might hangup through frustration, or to recover from an error, or by accident. The only sign of a hangup is that no input is received. This means that every routine that expects input must have a timeout, such that if no input occurs, the user is deemed to have hung up. This timeout adds a further complication, since the system should not hang up if the user is merely slow, so in every case where a timeout occurs, the system first asks the user to indicate her presence by striking any key. This is described by [5] as a channel checking dialogue. This dialogue can be a problem, too, since it introduces a timing condition in input: when the dialogue occurs the user may be uncertain which keystrokes count as part of the channel checking dialogue and which are considered to be part of the numeric input.

The Location Finder

The task of the Location Finder is to get the locations of the origin and destination of the traveler. Logically, the task is as shown in Figure 2. The actual interaction can be more complicated than this. Interaction complications arise from deficiencies in the program's database, inherent ambiguities in input, and user errors. This section discusses these problems, and concludes with an annotated transcript of the Location Finder, showing error recovery.

Handling Ignorance

MIT 02257

The first source of complication is that Direction Assistance is not omniscient. The database covers only a limited area of Boston, and is not complete even within that area. The user may enter an address outside the known area, or a phone number that is not in the phone book. In the former case, the user will probably hang up, since the program can't give directions to a place it doesn't know, but in the latter the user may be able to supply an address. So the first complication is the need to recover from a failure to comprehend a location,

The program has a few special case checks for phone numbers that are outside the local area. If the first digit of the number is 1, the program informs the user that the 1 is superfluous. The first three digits of a phone number specify the local office, and are sufficient to tell whether the number is outside the area covered. If the first three digits are not recognized, the program informs the user that the number isn't known, suggests the possibility that the user was entering an area code, and asks for the number again. This is useless, of course, if the user really intended a distant number, but helpful for typing errors.

Handling ambiguity

The second problem is that a street address may not be unique. The interface accepts only the name of a street, and not the type (e.g. "Street", "Avenue", "Place", etc.), so an entry such as "10 Beacon" is not unique. This problem can can also occur when the user enters a phone number, if the phone book entry does not include a street type. This situation is an example of partial incomprehensibility as described in [5]. The system informs the user of the problem, and asks for a sequential selection from the alternatives.

A street address can be made precise either by asking for the town or neighborhood (by place) or by asking for the type of the street. Some ambiguities are better resolves by place, and others by type. If there are three instances of a street, with two in one city, and the third in a second city, then asking for the city is not enough to resolve the reference fully. If the three streets are all of different type, it is better to ask for the type. The interface uses whichever will yield the most information, A sequential selection is employed,

Places can be named by city or by neighborhood. When resolving by place, the system always uses the least specific name for a locale that suffices distinguishes the street from others. Thus if two streets are different cities, the city name is used, but if both are in the same city, the neighborhood name is used. The neighborhood of a street is obtained by using the sip codes of the street as an index into a database of neighborhoods, organized as a tree.

The interface could have been designed to always collect this additional information, but that would have unposed a cost on every user. In addition, this method is faster, since the clarification dialog requires at most two sequential selections, instead of spelling both street type and city name.

User Errors

The most troublesome part of the Location Finder is the selection of the type of location to be entered (phone or street address). A user might forget this step, or might enter to wrong value. The interface handles all these cases. If the user skips the selection, then the next keystroke will be part of either a phone number or a street address number. If the number doesn't begin with 1 or 2, the selection routine will detect the error. If it does, than it will appear to be a valid selection, and the system will then prompt the user for one of these two things. The prompt itself may be enough to inform the user of the error. II not, there are four cases:

The user is entering a phone number, and the first digit is 1. The system is expecting a phone number, because the user hit 1. The user begin to enter the phone number. After three digits have been read, the interface checks that the exchange is known to the system. If it isn't, an error message is issued, and the user must begin again.

The user is entering a phone number, and the first digit is 2. The system is now expecting an address number. The address number input routine assumes that street numbers in Boston never exceed four digits, so when the fifth digit is entered, an error message is issued,

MIT 02258

The user is entering a street address number, and the first digit is 1. The system is expecting a phone number. When the user strikes number-sign to delimit the number, the phone number routine complains with the message "I thought you were entering a phone number, but you just entered a number sign, which you use only for addresses. If it was an accident, you can continue with the phone number, otherwise, enter number-sign again, and then you can enter the address.".

The user is entering a street address number, and the first digit is 2. Effectively, the first digit of the street address number is lost. The system does not detect this error. When the user strikes number sign, the system repeats the number. The user has a chance to abort the entry during the spelling of the street name, as described below.

This error checking is only possible because phone numbers and street address numbers have different syntax. A phone number is always seven digits, and the first three are chosen from a limited set. A street address number is always less than five digits, and is always delimited by a number-sign.

User Initiated Corrections

The user may detect an error while still entering data. The most frequently detected error is to find oneself in the wrong mode, that is, the user intends to enter an address, but the system expects a phone number, or vice versa. It is difficult to allow the user to correct mistakes, because there is no spare key to use for corrections. Star is reserved for help, and number-sign is sometimes is used for a delimiter (for numbers and names). The best solution is to use different abort conventions for the different types of input. The number sign key is always an abort key for phone numbers, and the zero is always an abort for names. For street numbers one aborts by entering star twice. This complication seems to be only way to provide the needed features.

Since the error correction mechanisms are not elegant, the initial prompts do not describe them. Instead, they are described by the "help" routines. If the user does not know how to correct an error, she will either hit "help", or will pause long enough for a timeout to occur, and this in turn will remind her to hit "help".

If the user makes a mistake that is not syntactically invalid (enters the wrong phone number, spells a differ- ent street) there is no way for the system to detect the error. Nor does the system confirm the location with the user. It was felt that adding a confirmation step (a yes o no question) would be too troublesome for the user. (In retrospect, this was a mistake.)

An annotated transcript

What follows is an annotated transcript of a dialog with the Location Finder. This dialog shows a few of the error recovery mechanisms in the interface. Audio output is shown in `this` font, keypad input like **this**, and comments are in *italics*

`Tell me where you are now. You can give me either a telephone number or an address. If you know the number of your phone, and it is not a payphone. then enter 1. Otherwise enter 2 and I'll give you further instructions`	*This is an example of an enumerated selection*
1	
`Enter the phone number`	
867	*The user realizes she's made a mistake, but doesn't know what to do. A long pause ensues, and the input times out.*
`For help enter *`	*The first layer timeout handler simply reminds the user how to get help.*
*****	
`You are entering a phone number for your origin. So far you've entered 867. If you make a mistake, or if you don't want to enter a phone number, enter number sign.`	MIT 02259

#

`Enter 1 for a phone number, 2 for an address.`

This restarts the Location Finder.

The system uses a shorter prompt the second time. This prompt also serves as implicit confirmation of the abort from entering a phone number.

1

`Enter the phone number`

8761111

`876-1111. 245 First Street, Cambridge.`

Confirms the entry and gives the location.

`Tell me where you want to go. Enter 1 for a phone number, 2 for an address.`

Shorter prompt.

3

The user intends to enter an address but forgets to preface it with 2.

`Wait! If you want to enter a phone number, you must first dial 1. If you want to enter an address, you must dial 2.`

The enumerated selection routine detects this error, unless the number begins 1 or 8, in which case other handlers detect the error.

2

`Enter the number of the address, followed by number sign.`

33

The user forgets to enter a closing delimiter. Pause and timeout.

`So far the number is 33. If you're through, enter number sign. Otherwise, please continue. If you make a mistake enter star.`

#

`33`

Echoes the street number

`Spell the name of the street, followed by another number-sign. Spell the entire name, but don't include words like Avenue or Street.`

33462676

`33 Edinboro St, Boston.`

Only one name is "spelled" that way, and only street in the area is named `Edinboro". No disambiguation is required.

The Narrator

Recall that the output of the Describer is a tour, a sequence of actions to make to get from start to finish along the route. The task of the Narrator is to read the lines of the tour, one at a time, while the user writes them down. It is necessary that the system be able to repeat lines and to go back to the beginning of the directions. After each line, the user makes an enumerated selection for one of three actions: repeat the line, move to the next, or start over. This is the most important use of enumerated selection in the interface, since this selection is made many times per session.

Difficulties of Speech Synthesis

Synthesis of speech by rule is essential for this application, because the text to be spoken is not known ahead of time, But synthesis by rule has drawbacks. Synthetic speech is harder to understand than recorded natural speech, and understand synthetic speech imposes extra more cognitive load on the listener [7,8]. Text to speech rules mispronounce some words. This section examines these problems and the solutions used here. This interface used a Digital Equipment Corporation Dectalk, but all available synthesizers have similar problems.

MIT 02260

The problem of intelligibility is most acute when reading directions. Names must be decoded on an acoustic phonetic basis alone. Neither syntax, semantics, nor pragmatics provide any constraint on names. The names of streets are usually unfamiliar to the user. Distances are similarly unconstrained. The Narrator copes with this by spelling all names in a line whenever the user asks the line to be repeated.

The Dectalk's pronunciation rules are good but not infallible for ordinary English. Perhaps a dozen words required either explicit phonetic pronunciations or explicit stress. For example, in this application, the word "address" is a noun, not a verb. A more annoying problem is that the rules are quite poor on proper names. Spiegel found that about 29% of surnames were pronounced incorrectly 112]. 1 found a similar result: The Dectalk's pronunciation was unacceptable for 220 of the approximately 1000 names. For these words, the interface uses the Dectalk's exceptional pronunciation dictionary facility. Ken Church has suggested that pronunciation of proper nouns might be improved by attention to etymology [1].

Prosodics

Prosodics refers to variations in pitch and timing of words. It can be difficult to obtain natural prosody with the Dectalk. Errors of prosody may simply sound unnatural, though comprehendable, or they may introduce great confusion into the interaction. Prosodics affect both the lexical content and the state of the entire interaction.

It is a fact of English intonation that the final word in a compound street name is stressed (e.g. First `Avenue, First `Boulevard) unless the word is "Street". Lists of letters also require special care. Since isolated letters are easily confused, they must be read slowly. Pitch is adjusted up at the beginning of the list, and downward just before the last element of the list.

Rhetorical constructions also required extra work on their prosodics. The error message given above ("If you want to enter a phone number, you must first dial 1. If you want to enter an address, you must dial 2.") uses a parallel construction. To be most easily understood, it should be spoken with accent on the digits.

Prosodics are also used to indicate the state of the discourse. The error message just mentioned is prefaced by an exclamatory "Wait!". This message is an interruption, not a response, so the emphatic tone produced by the exclamation mark tells the user that something unusual has happened, that something is wrong. In this case the Dectalk's prosodics were just what was needed.

A second discourse function of prosodics is the marking of turn endings. In this application, the interface is the only party speaking, so the issue of turn endings reduces to informing the user when input is required and also when input is not required. In natural conversation people use a mixture of cues from syntax, pragmatics, and prosodics to recognize the end of a speaker's turn, and the opportunity or obligation to reply [9] [10]. The limited prosodics of synthetic speech deprive the interface of one powerful method of indicating the state of the conversation. This leads to a problem most obvious when the system asks a yes or no question. A typical question, e.g. "If you want to enter a different address, hit any key now. Otherwise I'll ask for a new street name." is best expressed as two sentences, one for each alternative, but users often answer as soon as the first sentence is complete, lacking any cue that indicates that the turn is not finished.

Turn taking is even more of a problem for the Narrator. The interface requires the user to hit a key in order to hear the next set of instructions. Since an instruction may be several sentences long, the user has no syntactic clue as to when a chunk has ended. As a result, users do not always understand that the system is waiting for them instead of the other way around.

Further work on prosodics

Prosodic tuning is not difficult. The Dectalk provides a variety of means to affect the pitch and timing of utterances, including inserting explicit pauses between words or inserting stress or de-emphasis markers, or by altering the comma pause duration, or using the `pitch change" characters. The difficulty is simply that so much of it is required, and the means are not very powerful. The Dectalk does allow the programmer to explicitly designate duration and pitch for a

MIT 02261

phoneme, but this facility is impractical, since any attempt to do so imposes a monotone pitch on all unlabeled phonemes. It is not possible to supply a pitch contour for only a portion of the utterance

It is not surprising that the Dectalk's rules fail to provide natural prosody, since its production would require full comprehension of the matter to be synthesized, and this is clearly not to be expected. What would be a reasonable next step for products like the Dectalk, or systems that use them, would be to allow the programmer to explicitly mark the utterance for, e.g. rhetorical construction, and to employ intonational rules to synthesize the correct pitch contours for the desired effect. Research is required to discover both useful representations for this information and rules for synthesizing the correct intonation given the required marking. Such work is underway here at the Media Lab and also at Bell Labs [6,2].

A second main area of research is to provide a more natural method for listener control of speaker output. When people listen to spoken directions they typically use much more subtle means to let the speaker know how fast to go, or when to pause - they make inarticulate noises which we usually notate as "uh-huh" or "mm-hmm", sounds we call paraverbals or back channels 114]. A program capable of recognizing and responding to listener's paraverbals would be much more natural than one which used keypad presses. This is now a subject for research at the Media Lab[11]

Conclusion

Synthetic speech and key pad input are sufficient for construction of a robust and easy to use interface that can gather a wide range of types of data input. Research into synthesis of prosody by rule will make programming interactive systems easier. Until such systems are available, speech synthesizers should allow programmers to specify prosody explicitly. Recognition of listener paraverbals will make more natural interfaces possible.

Acknowledgments

Direction Assistance was written by the author and Thomas Trobaugh during the summer of 1985. I am grateful to those who tested the interface and made useful suggestions, and to Thinking Machines Corporation for supporting this work. Further development has occurred at the Speech Research Group at the Media Lab. At this writing, one version of Direction Assistance is part of the `Age of Intelligent Machines" exhibit touring science museums across the United States, and another is installed at the Computer Museum in Boston. These installations differ from the one described in this paper in that they do not have a telephone number to address database.

Footnotes

*1 This is a revised and expanded version of a paper that appeared in the Proceedings of the American Voice I/O Society conference of September 1986. The interface has changed, hopefully for the better, since the time this paper was written. This memo also supercedes an earlier memo titled "Giving Directions". The principle changes have been to add references to newer work.

*2 In newer versions of the program, Q is on the 7 key, with PRS, and Z on the 9 key, with WXY.

*3 One possible solution is to employ acoustic icons, as discussed in [4] The system might play a faint dial tone to indicate that a phone number was expected, and ambient street noise to indicate that an address was expected.

Bibliography

[1] Ken Church. Stress assignment in letter to sound rules for speech synthesis. In *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*, pages 246-253, July 1985.

[2] James R. Davis and Julia Hirschberg. Automatic generation of prosodic support for discourse structure. In

Proceedings of the Association for Computational Linguistics , page (submitted), 1988.

[3] James R. Davis and Thomas F. Trobaugh. *Direction Assistance* Technical Report 1, MIT Media Laboratory Speech Group, Dec 1986.

[4] William W. Gayer. Auditory icons: using sound in computer interfaces. *Human Computer Interaction* 2(2):168-177, 1986.

[5] Phillip J. Hayes and Raj Reddy. Steps towards graceful interaction in spoken and written man-machine communication. *Int I. Man-Machine Studies*, 19:231-284, 1983.

[6] Julia Hirschberg and Janet Pierrehumbert. The intonational structure of discourse. In *Proceedings of the Association for Computational Linguistics* pages 136-144, July 1986.

[7] Paul A. Luce, Timothy C. Feustel, and David B. Pisoni. Capacity demands in short-term memory for synthetic and natural speech. *Human Factors*, 25(1):17-32, 1983.

[8] David B. Pisoni, Howard C. Nusbaum, and Beth 0. Greene. Perception of synthetic speech generated by rule. *Proceedings of the IEEE*, 73(11):1665-1676, Nov 1985.

[9] Harvey Sacks, Emanuel A. Schegloff, and Gail Jefferson. A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4):696-735, 1974. Reprinted in *Studies in the Organization of Conversational Interaction*, J. Schenken, ed., Academic Press 1978.

[10] Deborah Schaffer. The role of intonation as a cue to turn taking in conversation. *Journal of Phonetics*, 11:243-257, 1983.

[11] Chris Schmandt. Employing voice back channels to facilitate audio document retrieval. In *Proceedings ACM Conference on Office Information Systems*, 1988. (to appear)

[12] Murray F. Spiegel. Pronouncing surnames automatically. In *Proceedings of 1985 Conference, American Voice I/O Society*, Sept 1985.

[13] Ian H. Witten. *Principles of Computer Speech*. Academic Press, 1982.

[14] Victor H. Yngve. On getting a word in edgewise. In *Papers from the Sixth Regional Meeting*, pages 567- 578, Chicago Linguistics Society, 1970.

This document was produced by scanning in the paper original, then manually adding HTML tags. I regret that I am unable to include the figures at this time. Apologies for errors introduced in the process - JRD, 4 Nov 1998.

EXHIBIT

23

Direction Assistance

James Raymond Davis
and
Thomas Frank Trobaugh

December 1986

Speech Research Group Technical Memo 1
The Media Laboratory
Massachusetts Institute of Technology

Abstract

Direction Assistance is an interactive program that provides spoken directions for automobile travel within the Boston area. The program has a telephone interface which uses touch tone keypad input and synthetic speech output. Routes are both short and easily followed. The directions are given in fluent English. The program has successfully directed newcomers through Boston.

This paper tells how we built Direction Assistance, with emphasis on how the available databases are and are not useful for this application. It also talks about automatic generation of route descriptions, and compares our work to that of others.

1 Introduction

1.1 Overview

Direction Assistance consists of about 11,000 lines of CommonLisp code, runs on a Symbolics Lisp Machine, and uses a Digital Equipment Corporation DecTalk synthesizer. It was written mostly during the summer of 1985 at the Thinking Machines Corporation of Cambridge, Mass. Since then, it has undergone periodic rewrites. It is running at the Media Lab, and is also installed at the Computer Museum in Boston and as part of the Age of Intelligent Machines exhibit traveling across the United States.

Direction Assistance consists of five modules. The Location Finder queries the user to obtain the origin and destination of the route. A location may be specified as a street address or as a telephone number. The Route Finder finds a simple, short route between the two points. The Describer generates high quality English text describing the route. The Narrator recites the route to the user. In addition, there is a graphical interface for maintenance.

These modules share a set of databases. The most important is the street map, which covers an eleven square mile area of Boston centered on the Charles River. A second database is an inverted phone directory, which provides a street address for a phone number.

In this paper, we discuss the databases, the Route Finder, and the Describer. The Location Finder and Narrator are described in [2].

It would be inappropriate to continue without mentioning the pioneering work of Jane Elliot and Mike Lesk[5, 4] . Our work differs from theirs in several ways. Our interface uses synthetic speech and pushbutton telephones rather than a graphics terminal. We are much more concerned with generating fluent English text than they. On the other hand, we are not much concerned with route finding algorithms. Finally, Elliot and Lesk used a Yellow Pages database in

MIT 02264

addition to the white pages and street map. We will not clutter this paper with citations to Elliot and Lesk on every point where they have made contributions. They are to be assumed.

(Figure Cartoon from The New Yorker, April 6 1987 p 56. A man sits by the side of a road with a sign reading: Lost? Pull over and let me draw you a map. $1.00")

We next discuss the underlying databases, and then the modules which use them. The description of the databases will by necessity refer to features of the program in order to motivate the construction of the database.

2 Databases

2.1 Streets

Our street map began as a DIME (Dual Independent Map Encoding) file distributed by the United States Geological Survey[1]. A DIME file consists of a set of straight line segments, each with a name, a type, endpoints in longitude and latitude, and some additional information. Segment types include natural features (chiefly water boundaries), railroads, town and property lines as well as streets. The latter are also labeled with address numbers on both sides of the street at each endpoint; thus is is possible to estimate the coordinates for any street address by interpolation, assuming all lot sizes to be constant.

We began with an 11 square mile subset centered roughly on the Charles River. This includes portions of Boston (Charlestown, Allston, Back Bay, South End, North End), Brookline, and Cambridge (Cam- bridgeport and Harvard, Inman, Central and Kendall Squares). (See figure 1.) There are about 279 miles of streets in the map, which contains 6163 segments, of which 5506 correspond to streets. The total size is about 477 kilobytes.

Figure 1: Street Database

The DIME file as supplied was far from suitable for our use. It contained many errors: streets were missing, mislabeled, or misconnected, and names were not spelled consistently. In some cases, more than one segment occupied the same place, and some segments were connected to themselves. We wrote a battery of plausibility checkers to detect and remove these errors, automatically where possible.

In addition to correcting errors, we had to add new kinds of information to the database. The most important information was whether a street was one way. We also classified streets by quality, and recorded textual descriptions for some turns. We'll now describe each of these.

Segments in the DIME file are deemed to connect if they share a common endpoint. We refer to this kind of connection as physical connectivity. Every segment has two endpoints, and for each of these there is a list of the segments which are physically connected to that endpoint. Obviously, physical connectivity is a symmetric non-reflexive relation. Physical connectivity is not sufficient for route finding, since it may not be legal to drive from one piece of pavement to another, even though they meet, because one might be one-way, or a turn might be forbidden, or there might be a divider in the way[*1]. To provide for the fact that one can not always drive from a segment to any other physically connected to it, we added a second kind of connection, legal connectivity. Two (street) segments are legally connected if one may drive from one segment to the other without breaking a law. Legal connectivity supplements, but does not replace, physical connectivity. Physically connected segments include those that can be seen in passing, and must also be retained, for they are important in forming descriptions. One cannot turn onto a railway, though the street and railroad segments are physically connected, but one may also wish to mention the crossing of the railroad as a salient detail of the tour.

Not all streets are created equal. We wanted our routes to use the widest, fastest, and most easily located streets, so we gave each street a value for goodness (super, good, average, or bad). By definition, most streets are average. The super streets are the expressways, interstate highways, and other limited access roads. Our rating of super is awarded more on the basis of being easy to find and to follow, since super roads are often crowded and slow. At the other extreme, the

MIT 02265

bad streets are those we know to be narrow or in poor repair. Our database contains only three miles of such streets. Unlike the taxi driver, we are not interested in shortcuts which use marginal streets.

The concept of "better than average" is a bit hard to define. We wanted to identify streets which were likely to be easy to find and follow. We decided that streets that were long were likely to be important, so we marked all streets longer than one half mile as "good", and then added a few more by hand if they seemed important. The resulting network is about 105 miles long, and forms a simplified skeleton covering our map. It appears in figure 2.

The third extension was to expand the street classification scheme. We added new segment types for bridges, underpasses, rotaries, and access ramps. This information is useful to both the route finder and to the describer, as we show below.

Finally, at every intersection in the map we can store additional descriptive information about each possible turn at the intersection, in the form of labelled items. Each item has a label telling what kind of information is stored, for instance an exit number or the text of a sign present at that intersection. This information is used by the Describer.

We made almost all of these corrections and augmentations ourselves from observations in the field. We could not find a paper map listing all the one way and restricted turning streets of Boston, so we had to drive around looking for them. This investment in time and effort is a. major cost of the system, but needs to be done only once. The graphic database editor was extremely useful, as it permitted rapid editing of the database. We commend the many designers of the Lisp Machine window system for making this easy.

Figure 2: Network of good streets

2.2 Neighborhoods

A related database lists the neighborhoods of Boston, with their associated zip codes. We need this database because a given street might occur in several different towns. For instance, there are three distinct streets named "Washington" in our map, in Boston, Cambridge, and Somerville. Even worse, Cambridge contains two different streets named "Elm".

The Location Finder uses this database to disambiguate street names. When the user supplies a name that could designate more than one street, it is necessary to ask for further information, e.g. "Do you mean Beacon Street in Cambridge or in Boston?". To make this as easy as possible, it is best to use the names of the most general locations that still distinguish the streets [*2]. If the street occurs in two neighborhoods of the same city, the neighborhood name is used. If the street occurs in different cities, the city name is sufficient. We determine neighborhood from the Zip code of the street. The mapping from Zip to neighborhood is imperfect, but good enough for our purposes. For the most part, the neighborhood names are those used by the local post offices. We think it is very likely that these names are also familiar to the local residents, and intelligible to visitors, but we have no evidence.

Figure 3: Central Square, 02139

2.3 Inverted Phonebook

The inverted telephone directory allows us to map telephone numbers to street addresses. We built this database ourselves, by inverting a "white pages" database. This required parsing the street addresses in the white pages. which was difficult for several reasons. The white pages have a great variety of spelling and abbreviation. We found, for instance, 23 variations of "Massachusetts". In addition, the format is not consistent. Sometimes listings contain professions ("atty" or "archt"), or a second phone number ("If No Answer"), or other information (e.g. "toll free", "children's phone"). We did not have the typographic information that helps separate names from locations and phone numbers. Finally, addresses are often incomplete, listing only a city, or road, or some a name which does not correspond to a street, such as a shopping center or an office park.

Even after parsing, it can be hard to determine locations from a phone book listing. Even the best entries have at best a

street, number, and city. But as we said above, streets occur in more than one place within a city. There is a rough correspondence between exchange and locale, so we can sometimes determine a unique location with this extra information. But when we can not, the Location Finder must ask the user to choose a location, as it does for street names.

Having described the databases, we now turn to the modules of Direction Assistance.

3 Route Finder

The Route Finder finds a route subject to three constraints. The route must be easy to follow, reasonably short, and it must be found before the user loses patience[*3]. These constraints conflict. Rarely is there a straight line route - the shortest route may require devious shortcuts. We are biased towards simplicity, since we want our users not to get lost.

The output of the Route Finder is a path, an ordered list of street segments, such that the origin is on the first segment, the destination on the last, and each segment is legally connected to the next. The real time requirements of the system rule out exhaustive, breadth first search[*4]. The current implementation uses a best first search that provides reasonably good routes in a moderate time. A sample route appears in figure 4.

Figure 4: A sample route

Best first search is an improvement on breadth first search. Search is conducted in (simulated) parallel on a list of candidate partial paths. For each path, there is a cost which is the sum of the known cost for the current path and an estimate which is a lower bound on the cost for (as yet undetermined) remainder of the path. At each step of the search, we consider the path of least cost, and expand it by considering all segments legally connected to its terminal end. The estimation function is just the Cartesian distance, since no route can be shorter than a straight line. Figure 5 shows every segment visited by the search in finding the route shown above.

Figure 5: All segments touched by search

As Elliot and Lesk point out, it is not desirable to find minimum distance routes, for these have too many turns. Such routes are hard to describe and hard to follow. Elliot and Lesk impose a cost of 1/8 mile for a right turn, and 1/4 mile for a left turn. We extend their system of costs in several ways. First, we consider street goodness. Travel down a "super" street is not as "expensive" as travel down an average street, and travel down a "bad" street incurs a surcharge. Second, we consider sharp right turns to be as bad as left turns, since they are harder to execute. Third, we reduce or waive turn costs in some cases. For example the turning cost is halved for a turn on to or off a one-way street, and waived altogether for a forced turn ("left turn only"). A turn onto a bridge is also free, since bridges are major landmarks, and contribute to ease of following the route. We have not studied the effect of these routes on the routes found, nor have we attempted to determine whether the routes are better where different. Such a study would require a model of driver's errors, both of understanding and of execution.

4 Describer

The Describer generates a set of text instructions for following the route. (An example of its output appears in figure 6.) We generate text instead of a map for two reasons. First, the system is used by telephone, which limits the output to voice. But even if our users had portable graphics terminals with modems, we would prefer text to graphics, because some people can not read maps. In a survey of map reading abilities Streeter and Vitello recommend text as a "lowest common denominator" [9].

The Describer creates a new representation of the route, instead of using the path itself. There are two reasons for this second form of representation. First, the elements of a a path (segments) are too fine grained for useful textual description. Recall that a segment reaches from just one intersection to the next. This is smaller than our sense of a "street", which continues as a unity past many intersections. In addition, segments are straight lines: so a street with no intersections might be still represented as a sequence of segments if it made a broad turn. We want to describe the

entire stretch of a street as a single object. A second reason is that a path is just a topological structure, but natural instructions should be expressed in terms of geometry and of types of streets. Consider the difference between a "fork", a and an "exit", as shown in figure 7. All have the same topology - a branch in the road. But they must be described differently. The Describer's structure is a tour, which is a sequence of acts to be taken in following the path.

> If your car is on the same side of the street as 20 Ames Street, turn around, and start driving. Drive all the way to the end, about one eighth of a mile. Make a left onto Memorial Drive. Drive about one eighth of a mile. After you pass Wadsworth Street on the left, take the next left. It's an easy left. Merge with Main Street. Stay on Main Street for about ninety yards, and cross the Longfellow Bridge. You'll come to a rotary. Go about half way around it, and turn onto Cambridge Street. Drive all the way to the end, about three quarters of a mile. Make a right onto Tremont Street. Drive about one half of a mile. After you pass Avery Street on the left, take the next left onto Boylston Street. Stay on Boylston Street for about one eighth of a mile. After you cross Washington Street, it becomes Essex Street. Keep going. Drive about one eighth of a mile. After you pass Ping On Street on the right, take the next right onto Edinboro Street. Number 33 is about one eighth of a mile down on your right side.

Figure 6: sample of directions

Figure 7: T, fork, and exit all have same topology

4.1 Acts

Acts are things a driver does (or notices) while following a route. Figure 8 shows our taxonomy of acts.

- Boundaries
  - Start
  - Stop
- Straight
  - Name Change
- Turn
  - Enter
  - Exit
  - Merge
  - Fork
  - UTurn
  - Rotary
  - Ordinary

Figure 8: Act Taxonomy

Each of these acts must be recognized. The route finder works only with segments, and the Describer builds acts which describe motion from segment to segment. We now describe each of these acts, and how we recognize them. We describe the text generated for each below.

The first act is necessarily START, and the last STOP. They are trivial to recognize. The NAME CHANGE act requires the driver to notice a change in name, but nothing further. We include it only to avoid confusion. The difference betweeb a NAME CHANGE and a TURN is that the former consists of a two streets meet within 10 degrees of straight, and where there is no other segment at the intersection with the same name as either of them. These two criterion are almost correct, but not quite right. There are streets which seem (to us) to be name changes, but have more extreme turns (at least, as represented in the map). For the present, we have caused these to be treated as name changes by changing the map, slightly altering the positions to make the turns more gentle. This would be intolerable were we using the map for, say, surveying, but is of no consequence for route description.

MIT 02268

There are several types of TURN acts. The ENTER and EXIT acts refer to limited access roads. In this case, some of the travel will often be on "nameless" segments - access ramps. This shows one reason for the additional classification of street segments. We want to recognize entrances and exits, and we want to describe access ramps in different terms than other streets.

A MERGE and a FORK are similar in that they are different actions that might be taken at the same intersection, depending upon the direction one is driving. A Merge has the following characteristics:

1. Old and new streets have different names.
2. Only one street is legally possible.
3. The angle of turning is small.
4. There is at least one other street going to the destination street.
5. All streets make only small turns onto the destination.

At a FORK on the other hand, there are at least two ways to go, though all are shallow turns. Note that a "fork" onto an exit ramp is recognized as an EXIT.

There are two types of U turn known to drivers in Boston. The first kind is made in the middle of the street (within a single segment, in our representation). Our routes never include such turns. Not only are they illegal, such moves never shorten the path. The second kind of U turn is the sort one makes to reverse direction on a divided road. Typically one makes a left onto a nameless piece of road, which is often very short, and then makes a second left. This double turn is what we call a U TURN act. It is very important to recognize this act, because describing it as two successive lefts is very confusing. It is a single entity in the minds of drivers. We recognize a U TURN as a pair of turns where the intermediate segment is less than 165 yards long, the total angle is within 20 degrees of 180, and the name of the street is unchanged after the two turns.

Perhaps the most insidious feature of Boston's streets is the ROTARY. For those not familiar with the term, a rotary is a one way street in a circle. Traffic enters the rotary on roads which are (usually) tangent to the circle, moves counterclockwise around the circumference, and exits on another tangent. Rotaries are difficult to traverse because they cars enter and exit within a very short distance, without much room to maneuver. Recognition of a rotary is trivial, but only because we label all rotary segments explicitly in the street map.

An ORDINARY turn is anything not handled by one of the above cases.

4.2 Cues

While the Describer is collecting the acts of the tour, it also collects cues. A cue helps the driver follow the tour. We distinguish four kinds of cues. Action cues tell when to do an act. Confirmatory cues describe things that will be seen while following the route. Warning cues caution the driver about possible mistakes. A warning successfully heeded also serves as a confirmatory cue. Failure cues describe the consequences of missing an act, e.g. "If you see this, you have gone too far".

The most common action cue is just the name of the street. An instruction such as "Turn right onto Tremont Street." tells the driver what to do and when to do it. This cue may be hard to follow, since street signs may be missing. A very strong action cue is coming to the end of a road. No one is likely to forget to turn under this circumstance, since the alternative is to leave the road. We refer to this as a "forced turn" cue.

Distance traveled is also a cue, but hard to use. People have a vague sense of distance, but not an accurate one. Still, we use distance as a secondary cue, because we can compute it easily and it helps some people. We express distance in yards when less than 1/16 of a mile, and other distances in approximate fractions of a mile because people are accustomed to seeing distances expressed this way. We do not use tenths of miles, because some people do not know how to use odometers, and because using an odometer to calculate distance requires doing mental arithmetic, which might prove distracting while driving.

We never use blocks, since a block is not a clearly defined concept. We do not know whether a block is bounded by an intersecting street, or only by streets that cross and continue. Figure 9 illustrates this. In any event, we do not expect our drivers to be able to drive more than two or three blocks without losing count. Since we don't want to rely on distance or counting blocks, we use as a cue for an act the name of a street immediately preceding the act. This is a risky cue, since the driver who misses the cue may keep looking for it and miss the destination street as well. To make this less likely, we use only streets on the same side as the turn for a cue. This way, a driver need attend to only one side of the road while looking for street signs, so if the cue street is missed, the target street may still be seen. This same strategy is adopted in [10].

Figure 9: Is the distance between "A" and "B" one block, or two?

The confirmatory cues are crossing major streets or railroads, or going through an underpass. The only warning cue currently is a warning about left exits from limited access roads. We assume drivers will not take the wrong exit, but if they keep in the left lane they may be surprised by an unexpected left exit. We have not implemented failure cues.

4.3 Generating Text

For each act there is a corresponding routine which generates one to three sentences describing it. The routine selects appropriate cues from all those gathered. Now we'll describe some aspects of generating text.

```
(defun disc-seg-rotary (act)
   (list
    (make-sentence
      "You'll" "come" "to"
      (make-np-constituent `("rotary")  : article indefinite))
   (make-c onj unction-sentence
     (make-sentence
       "Go" (rotary-angle-amount (get-info act `rotary-angle))
       "way" "around" (make-anaphora nil "it"))
     (make-sentence
       "turn" "onto" (make-street-constituent (move-to-segment act) act)))))

(defun rotary-angle (angle)
   (selector angle <
    (45 `("just" "a" "little"))
    (135 `("about" "a" "quarter"))
    (225 `("about" "half"))
    (315 `("about" "three" "quarters" "of" "the"))
    (360 `("almost" "all" "the"))))
```

Figure 10: Generator for rotary

Generating text for a START is tricky because it is hard to specify an initial direction. We do not use absolute directions, because most people do not know them. If we had a landmark database we might sometimes use relative direction (e.g. "towards the river"). Instead, we use the initial address, since that also determines a side of the street, and thus a direction to drive. We might have used "If your car is on the same side of the street as ... start driving the way it is facing.", but that sounded clumsy. Instead, we chose to give a negative instruction, either "If your car is on the same/opposite side of the street from turn around, and start driving." For one way streets we mention that the street is one way, and say "just start driving." We think (but do not know) that drivers would not have confidence in the instruction ("just start driving.") if it did not indicate that the system knew about the one way street.

One of the simplest generators is for rotaries. It appears in figure 10. Rotaries are hard to describe and hard to follow, because there are no good references for distance around a rotary. We can not expect people to measure angular distance around the rotary, and there may not be signs. The segments of a rotaries may or may not be nameless, or there may be several names involved. The rotary itself may have a name, e.g. Leverett Circle, but this name does not appear in the database and usually does not appear on any street signs either.

MIT 02270

Output from this generator appears in figure 6. The generator produces two sentences, the second of which is a conjunction of two sentences. The distance around the rotary is converterted from an absolute angle, as measured on the map, to an approximation in English.

The instructions generated have syntactic structure only for sake of exploiting generality in text generation. Thus the function make-np-constituent handles agreement between the article and the noun. The function make - sentence ensures that capitalization and punctuation are correct. Text is sent directly to the synthesizer, and punctuation is required to achieve proper intonation. The function make-anaphora serves no purpose at present, but in planned future research will allow us to convey intonational features of discourse[3].

4.4 Comparison

We can compare our descriptions with those generated by Streeter and colleagues[10].

Streeter's descriptions are intended to be understood and acted upon in real time, as if uttered by a navigator in the next seat. (In fact, they are recorded on a tape, and the driver pushes a button to play the next instruction.) This interface imposes a new requirement on the form of the directions. Since they are to be heard and acted on in real time, it is important to repeat essential information so that it can be remembered. In our interface, we assume people are writing down the directions before they begin to drive, so repetition is not crucial. (The user can ask the Narrator to replay an instruction if it is not understood.)

They classify turns into ordinary turns, "T" turns, complex intersections, turns in short succession, and continues. Their "T" turn is our "forced turn" cue. The difference between an ordinary turn and a "T" turn is that the latter needs no failure cue. So our treatments are similar. We do not distinguish complex intersections, though we should. The Route Finder should avoid them, and the Describer should warn about them.

Their instructions are sometimes more structured than ours. They cluster turns which occur close to each other into a single instruction block, and their "continue" is just our "name change", but is also incorporated into the following turn. We recognize the importance of providing higher levels of structure, and wish to remind the reader that Streeter and company were working by hand, not with a program, and were in a better position to form hierarchies than we.

We claim that our directions are more natural than those of Elliot and Lesk, but have no proof for this. We leave it to the reader to judge.

Are our directions clear? We know that people have been able to follow them, but we have not made any systematic test. Christoper Riesbeck wrote a program (McMAP) which judged the clarity of directions. Our directions would not be acceptable to it, to judge from its published description. Partly this is because we talk about features the program does not know, for example rotaries, but also because his program explicitly rejects use of miles for distance as inherently unclear. We use mileage only as an approximation, as a cue for when to look for a landmark, but the weak syntactic powers of McMAP would not notice this. Also, we use "drive all the way to end", which Riesbeck terms a "procedural operator", and did not implement. Since people accept our directions, this suggests that Riesbeck's rules are too strict, or perhaps not powerful enough.

5 Discussion

Products like Direction Assistance are beginning to appear in the marketplace. It is reported that ETAK, of Sunnyvale California, has a product (the Navigator) which, installed in a car, estimates the car's position by counting wheel rotation and turning angle and comparing results with a stored map. A display in the dashboard displays the local area and the position of the car. The Navigator does not supply driving directions, but surely could be made to do so.

A more similar product is DriverGuide, made by Karlin and Collins, also of Sunnyvale, which is reported to produce printed directions for travel in the Bay Area[8].

5.1 Better databases are required

Any serious use of Direction Assistance requires further improvements to the street map. The area covered is too small, and even the small region covered is not fully mapped. More significantly, there are additional facts that the current street database format can not represent.

Among these are time-dependent legal restrictions (e.g. "no left turn during rush hour"), restriction of height, weight, and prohibition of commercial vehicles, multiple names of streets, presence of stop lights, and landmarks. In addition, the representation of addresses is not sufficient. We have seen addresses with fractions and with letters, and there are also streets where both even and odd numbers are on the same side of the street.

A practical system must account for multiple names. When Route 93 passes through Boston, it is also Route 3, the Fitzgerald Expressway, and the Southeast Expressway. When Massachusetts Avenue turns north at Harvard Square, only the southbound lane is "really" Massachusetts Avenue. The other direction is officially Peabody Street. We do not know which name to use when naming these streets, but we should at least be able to accept all synonyms on input.

Boston, like any city, changes its configuration of streets daily. Some changes, e.g. for construction, are temporary, although they may persist for years. Others are permanent. Streets are built and removed, and sometimes they change names or directions. A practical system requires accurate and timely corrections to the database.

We could give better directions with a better database, giving, for example the location of traffic lights or landmarks such as gas stations. Elliot and Lesk were able to capture business locations from an online Yellow Pages. To be more ambitious, we might hope for a representation rich enough to capture the qualities of image and orientation described by Kevin Lynch[7]. We have no proposal for how to do this at present.

5.2 Applications

We initially designed Direction Assistance with tourists in mind. Boston's confusing streets often lead the visitor astray. A tourist's direction guide could be provided by the city, or as a profitable venture. But a tourist may not know the street address or phone number for the destination. In fact, there may not be one, for the destination might be a general area, such as a neighborhood or park. Tourists would probably prefer to identify locations by name. It might be difficult to add this feature without making the interface more complicated.

Direction Assistance could direct people to services. Given the caller's location and the type of service desired, Direction Assistance could select the closest, and provide a route. This service might be dedicated to a single vendor (e.g. for banking machines) or as an advertising service for many customers.

Routing delivery vehicles pose special problems. Some of the most useful routes in Boston are closed to commercial vehicles, either for legal reasons or because they have such low underpasses that even scofflaws can not get through. We could extend the street database to record such facts.

We also feel compelled to mention the implications of Direction Assistance for privacy. Should a public Direction Assistance include home telephone numbers? People may want to keep the ability to give out their home phone numbers without also revealing their addresses to callers. One can hang up on an annoying caller. A visitor may be harder to dispose of. Acknowledgments

A prototype version was written by Dinarte R. Morais during the winter of 1985. We are indebted to him for decoding the DIME files, the initial window system interface, and the proof of concept. We made extensive use of a database package and string matcher written by Craig Stanfill. Charles Lieserson made major improvements to the search algorithm of the Route Finder. Fletch McCellan of the PhoneBook Corporation loaned us the raw phone book database. This work would' not have happened without the guidance and persistence of Brewster Kahle. This paper was much clarified by the comments of Janet Cahn, Mike Hawley, Margaret Minksy, and Chris Schmandt. We thank them all.

This work was supported at MIT by the DARPA Space and Naval Warfare Systems Command, un- der contract numbers N00039-89-C-0406 and N00039-86-PRDXOO2 and by the Nippon Telegraph and Telephone Public Corporation. Hardware support was provided by Symbolics and Digital Equipment Corporation.

This paper bears the names of two authors, for the program was joint work. But though it is written in the plural, it is the work of only one of us. I dedicate it to Tom, who did not survive to see his work described. Though too small a memorial, it is the best I can manage at this time.

Footnotes

[*1] In this case, the turn is forbidden by physical obstacles, and not merely law or custom. But rather than engage in an epistemology of barriers, we use the same mechanism to represent this restriction.

[*2] This assumption could be tested. If people represent locales hierarchically, and if there is a preferred level of representation, it might be more difficult to determine inclusion in a too.general region.

[*3] A fourth constraint which we do not consider explicitly is that the route must be easy to describe. We are familiar with situations where a person asks for a route to a familiar place, but we can not describe the route because it is a "felt path": we no longer remember (or do not know) the names of the streets, only a list of subtle cues we can't describe.

[*4] on a serial machine, anyway. An experimental version on the Connection Machine [6] works in just this way.

References

[1] Geographic Base File GBDF/DIME: 1980 Technical Documentation. U.S. Department of Commerce, Data Users Services Division, 1980.

[2] James R. Davis. Giving directions: a voice interface to an urban navigation program. In *Proceedings of 1986 Conference*, pages 77-84, American Voice I/O Society, Sept 1986.

[3] James R. Davis and Julia Hirschberg. Automatic generation of prosodic support for discourse structure. In *Proceedings of the Association for Computational Linguistics*, page (submitted), 1988.

[4] R. J. Elliot and M. E. Lesk. *Let Your Fingers Do the Driving: Maps, Yellow Pages, and Shortest Path Algorithms*. Technical Report unpublished, Bell Laboratories, 1982.

[5] R. J. Elliot and M. E. Lesk. Route finding in street maps by computers and people. In *Proceedings of the National Conference on Artificial Intelligence*, pages 258-261, 1982.

[6] W. Daniel Hillis. *The Connection Machine*. MIT Press, 1985.

[7] Kevin Lynch. *The Image of the City*. MIT Press, 1960.

[8] Ronald Rosenberg. *Mapping out a new idea*. The Boston Globe, 39, 1987. February 17.

[9] Lynn A. Streeter and Diane Vitello. A profile of drivers' map reading abilities. *Human Factors*, 28:223-239, 1986.

[10] Lynn A. Streeter, Diane Vitello, and Susan A. Wonsiewicz. How to tell people where to go: comparing navigational aids. *International Journal of Man/Machine Systems*, 22(5):549-562, May 1985.

This document was produced by scanning in the paper original, then manually adding HTML tags. I regret that I am unable to include the figures at this time. Note that the printed version of this paper has the wrong year 1987, not 1986.

Apologies for errors introduced in the process - JRD, 4 Nov 1998.

EXHIBIT

24

FILED UNDER SEAL

EXHIBIT
25

FILED UNDER SEAL

EXHIBIT

26

IW 1415424

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

February 07, 2006

THIS IS TO CERTIFY THAT ANNEXED IS A TRUE COPY FROM THE RECORDS OF THIS OFFICE OF THE FILE WRAPPER AND CONTENTS OF:

APPLICATION NUMBER: ***07/565,274***
FILING DATE: ***August 09, 1990***
PATENT NUMBER: ***5,177,685***
ISSUE DATE: ***January 05, 1993***

By Authority of the
Under Secretary of Commerce for Intellectual Property
and Director of the United States Patent and Trademark Office

N. Williams

N. WILLIAMS
Certifying Officer

PART (2) OF (3) PART(S)

0001

5177685

07/563274 PATENT DATE JAN 05 1993 PATENT NUMBER

SERIAL NUMBER	FILING DATE	CLASS	SUBCLASS	GROUP ART UNIT	EXAMINER
07/563,274	08/07/90	364			

JAMES R. DAVIS, NORTH CAMBRIDGE, MA; CHRISTOPHER M. SCHMANDT, MILTON, MA.

CONTINUING DATA*******************
VERIFIED

NONE

FOREIGN/PCT APPLICATIONS**********
VERIFIED

NONE

CERTIFICATE
MAR 1 5 199?
OF CORRECTION

FOREIGN FILING LICENSE GRANTED 09/22/90

Foreign priority claimed / 35 USC 119 conditions met / Verified and Acknowledged	AS FILED	STATE OR COUNTRY	SHEETS DRWGS	TOTAL CLAIMS	INDEP CLAIMS	FILING FEE RECEIVED	ATTORNEY'S DOCKET NO.
☐ yes ☒ no / ☐ yes ☒ no / Examiner's initials	→	MA	2	58	1	$ 946.00	

SAM PASTERNACK
CHOATE, HALL & STEWART
EXCHANGE PLACE
53 STATE STREET
BOSTON, MA 02109

AUTOMOBILE NAVIGATION SYSTEM Using Real Time Spoken Driving Instructions

U.S. DEPT. of COMM. Pat. & TM Office—PTO-436L (rev. 10-78)

PARTS OF APPLICATION FILED SEPARATELY

NOTICE OF ALLOWANCE MAILED

PREPARED FOR ISSUE

Assistant Examiner | Docket Clerk

CLAIMS ALLOWED
Total Claims 58 | Print Claim

ISSUE FEE
Amount Due | Date Paid

Primary Examiner

DRAWING
Sheets Drwg. 3 | Figs. Drwg. 5 | Print Fig. 1

ISSUE CLASSIFICATION
Class 364 | Subclass 443

ISSUE BATCH NUMBER

Label Area

WARNING: The information disclosed herein may be restricted. Unauthorized disclosure may be prohibited by the United States Code Title 35, Sections 122, 181 and 368. Possession outside the U.S. Patent & Trademark Office is restricted to authorized employees and contractors only.

0002

Loc 036 #4 A/N

In the United States Patent and Trademark Office

Applicant : James Raymond Davis and Christopher M. Schmandt
Serial No. : 565,274
Filed : August 9, 1990
For : Automobile Navigation System

Examiner :
Art Unit :

Commissioner of Patents and Trademarks
Washington, D.C. 20231

Information Disclosure Statement

Sir:

Aspects of the invention have been described in the following sources which were incorporated by reference in the specification and are listed on the attached PTO–1449. Copies are enclosed.

1. "Synthetic speech for real time direction-giving," by C.M. Schmandt and J.R. Davis (*Digest of Technical Papers, International Conference on Consumer Electronics*, Rosemont, Illinois, June 6–9, 1989) is an abstract describing the goals of the research which resulted in the present invention.

2. "Synthetic speech for real time direction-giving," by C.M. Schmandt and J.R. Davis (*IEEE Transactions on Consumer Electronics*, 35(3):649–653, August 1989) is an expansion of the above abstract into a paper. (Kindly note that the publication date for this issue was September 8, 1989, as indicated in the accompanying copy of the certificate of copyright registration.)

3. "The Back Seat Driver: Real time spoken driving instructions," by J.R. Davis and C.M. Schmandt (*Proceedings of the IEEE Vehicle Navigation and Information Systems Conference*, Toronto, Canada, September 1989) describes the strategies employed by the present invention to successfully use speech.

4. "Back Seat Driver: Voice assisted automobile navigation," by J.R. Davis (Ph.D. thesis, Massachusetts Institute of Technology, September 1989) is the most detailed publication describing the present invention to date. The thesis includes a long list of references. Those deemed by the applicants relevant to the present invention as claimed are included on the enclosed PTO–1449 and are discussed below. If the examiner requires further information regarding any of the references cited in the thesis but not included in this Information Disclosure Statement, the applicants will be pleased to provide such.

A short news article on the invention appeared in the July 1990 issue of *Technology Review*. The article, entitled "Terminal Back Seat Driver," is listed on the attached PTO–1449 and a copy of the article is enclosed.

The following references were incorporated and discussed in the the specification. They are listed on the attached PTO–1449 and a copy of each is enclosed.

1. "CD-ROM Assisted Navigation Systems" by O. Ono, H. Ooe, and M. Sakamoto (*Digest of Technical Papers, IEEE International Conference on Consumer Electronics*, Rosemont, Illinois, June 8–10, 1988) describes the vehicle location system built by NEC Home Electronics, Ltd. which was employed in the working prototypes of the present invention. It is a dead-reckoning position keeping system which uses speed and direction sensors. To compensate for error, it uses map matching on a map database stored on CD-ROM.

2. "Attention, intentions, and the structure of discourse" by B.J. Grosz and C.L. Sidner (*Computational Linguistics*, 12(3):175–204, 1986) describes a discourse theory.

3. "The intonational structuring of discourse" by J. Hirschberg and J. Pierrehumbert (*Proceedings of the Association for Computational Linguistics*, 136–144, July 1986) describes a theory of intonational meaning.

The following references, listed on the attached PTO-1449, relate to further aspects of the present invention as embodied in the working protoype:

1. *Geographic Base File GBDF/DIME: 1980 Technical Documentation* (U.S. Department of Commerce, Data Users Services Division, 1980) describes the DIME format, the basis for the map database of the working prototypes of the present invention.

2. "A formal basis for the heuristic determination of minimum cost paths" by P.E. Hart et al. (*IEEE Transactions on SSC*, 4:100–107, 1968) is the first of many papers which discuss the A* search algorithm on which the route finder of the working prototypes of the present invention is based.

3. "Cellular telephone data communication system and method" by Harry M. O'Sullivan (U.S. Patent 4,697,281) describes the error correcting modem used in a working prototype of one embodiment of the present invention.

4. "A voice interface to a direction giving program" by James R. Davis (Technical Report 2, MIT Media Laboratory Speech Group, April 1988) describes the method of address entering using a cellular telephone keypad employed in one embodiment of the present invention.

The following discussion, summarized primarily from the Davis thesis cited above, examines prior automobile navigation systems. A related survey can be found in "Automated provision of navigation assistance to drivers" by Matthew McGranagha et al. (*The American Cartographer* 14(2):121-138, 1987). The references cited are included on the enclosed PTO-1449. The undersigned attorney does not currently possess copies of most of these articles, and therefore copies of most are not enclosed. However, at the request of the examiner, copies of any article cited will be obtained and forwarded to the examiner for consideration.

Early application of computers to navigation was intended to reduce traffic congestion by providing route information to drivers. In the Electronic Route Guidance System (ERGS) described in "An electronic route guidance system for highway vehicles," by D.A. Rosen et al. (*IEEE Transactions on Vehicular Technology* VT-19(1):143–152, Feb. 1970), a driver beginning a route finds the intersection closest to the destination, then enters a five letter code word for the intersection. When the vehicle passes over an induction loop sensor in the road, it transmits the destination to a central computer. The computer determines the best route, and relays instructions to the car. This interchange of information occurs at every instrumented intersection. Driving direc-

tions combine simple text from a nine word vocabulary and directional arrows, and are displayed on a "heads-up" display. The ERGS system was designed but never implemented. A similar system was designed and tested in Germany in the late seventies, and is described in "Function, Equipment, and Field Testing of a Route Guidance and Information System for Drivers (ALI)" by P. Braegas (*IEEE Transactions on Vehicular Technology*, VT-29(2):216-225, May 1980).

The pioneering work on computer navigation assistance is presented in "Let your fingers do the driving: Maps, yellow pages, and shortest path algorithms" by R.J. Elliot and M.E. Lesk (Technical Report, unpublished, Bell Laboratories, 1982) and "Route finding in street maps by computers and people" by R. J. Elliot and M.E. Lesk (*Proceedings of the National Conference on Artificial Intelligence*, pp. 258-261, 1982). Elliot and Lesk showed that specialized, as opposed to general purpose, graph search algorithms are desireable for route finding. One reason for this is that the shortest route may not be the best route. Elliot and Lesk added a system of weights, which are extra costs associated with turns. This caused the algorithm to prefer slightly longer routes with fewer turns to short, twisty routes. Elliot and Lesk also were the first to implement a program to generate written natural language driving instructions for the route. In their instructions, a route consists of a beginning, a sequence of turns and crossings (of rivers or railroads), and an ending. For each of these, there is a template comprising a sequence of words and slots, representing fixed and variable components of a sentence for a given type of act. The words are copied directly to the output, and the slots are filled in according to the particulars of an act. A third contribution of Elliot and Lesk was to integrate the digital map with other location oriented databases, including a services database and a personal address book. This allowed the program to find routes to addresses given a person's name, to find the closest store of a specified category, and to mention stores along the route as possible landmarks.

"Direction assistance" by J.R. Davis and T.F. Trobaugh (Technical Report 1, MIT Media Laboratory Speech Group, Dec. 1987) describes a system which provides spoken directions between locations in the Boston area. It uses a Dectalk speech synthesizer. This synthesizer includes a telephone interface, so it can answer a phone call and decode touch tone button presses. To use Direction Assistance, you call it from a touch tone phone. It answers the call, and prompts you to enter your origin and destination locations as street addresses. It finds a route, then describes the route to you. Direction Assistance was directly inspired by Elliot and Lesk, and extends their work in three ways. The most significant difference is that Direction Assistance speaks its directions, where Elliot and Lesk drew maps and provided written text. The disadvantage is that users must remember the instructions or write them down. A second significant difference between Direction Assistance and the work of Elliot and Lesk is that Direction Assistance generates better quality descriptions of the route. The improvement arises because the text generation process first analyzes the route into a sequence of "acts", and then generates descriptions from these acts, instead of working directly from the route. An act represents something that the driver does. There are eleven different acts, each representing a different way of moving. The type of act depends upon topology (how many streets are present at an intersection, and which way traffic can flow), geometry (what angles the streets make) and what kinds of streets are involved. There is also a

function to find an appropriate cue or landmark (e.g. a street crossed or an underpass) just before the location of the act. The Direction Assistance route finder uses a different algorithm (the A* search) than Elliot and Lesk, and has a different set of weights. The weighting scheme ranks roads by a four-valued "goodness" feature and penalizes routes that use less good roads by multiplying the mileage by a constant factor. It also reduces or waives the penalty for turning under a set of circumstance having to do with predicted ease of following; for example, a turn onto a one way street incurs a lesser penalty, since it is unlikely that the driver would turn the wrong way. It reduces the penality for "T" turns since the driver cannot possibly miss the place to turn. In practice, these weight changes have very little effect.

The Hertz car rental company offers "Computerized Driving Directions" at some of its rental counters. The directions include approximate mileage and estimated travel time, but are highly schematic, even cryptic. Despite appearances, these instructions are not computer generated. The Hertz system is more akin to a database retrieval system than a route finder. A California firm, Navigation Technology, sells a product called "DriverGuide" which is reported to be able to print driving directions between any two points in the San Francisco area ("Mapping out a new idea" by R. Rosenberg, *The Boston Globe*, February 17, 1987, p. 39 and "A Way to Go From A to B" by R. Alexander, *The New York Times*, March 11, 1989, p. 52).

Peeder Ma describes a system which gives textual directions in "An algorithm to generate verbal instructions for vehicle navigation using geographic database" (*The East Lakes Geographer*, 22:44-60, 1987). His work is similar to both Elliot and Lesk's and to Direction Assistance, but was apparently created independently of both. Ma uses A* search with a penalty factor to minimize the number of turns. Unlike Elliot and Lesk, he uses the same penalty for both left and right turns. His street map representation does not include one–way streets or restrictions on turning ("on left turn") so it does not always find usable routes. His route descriptions use a taxonomy about as elaborate as that of Direction Assistance, but the text generated is more stylized.

The systems of Elliot and Lesk, Ma, and Hertz provide static, textual instructions. Direction Assistance gives static verbal instructions. The limitations inherent in static navigation systems were discussed in the specification.

Several groups have built position or navigation systems for use *in* automobiles. For the most part, these systems have not been well described in the literature, probably from a desire to preserve commerical secrecy. The following discussion summarizes the applicants' knowledge of such systems, as gained through the cited articles.

The most well known automobile navigation system in the United States is the ETAK Navigator, which displays the car's position on a map display on the dashboard. The system is described in "Extending Low Cost Land Navigation Into Systems Information Distribution and Control" by S.K. Honey, M.S. White and W.B. Zavoli (*IEEE Position and Locations Symposium*, pp. 439-444, 1986, IEEE 86CH2365-5) and "Map Matching Augmented Dead Reckoning" by W.B. Zavoli and S.K. Honey (*Proceedings of the 35th IEEE Vehicular Technology Conference*, pp. 359-362, 1986, IEEE CH2308-5). The map rotates as the car turns so that the forward direction is always straight up on the map. The system provides a limited amount of navigation assistance. The driver may enter a destination (as a street address or intersection), and the system will dis-

play the direction and distance to the point. It remains the driver's task to select an appropriate route to the destination.

The Routerechner, described in "On board computer system for navigation, orientation, and route optimization" by P. Haeussermann (Technical Paper Series 840483, Society of Automotive Engineers, 1984), was designed to provide directions in and between German cities. The route finder could receive real time traffic information by digital radio while on route. This system's map included only the Autobahn, and not the cities (this was before CD-ROMs were widely available), yet it also provided a limited navigation service within cities. The user entered the destination as a pair of coordinates, and the system displayed the direction and distance to the destination. As with ETAK, it was the user's responsibility to select an appropriate road.

The Honda Electro Gyro-Cator, described in "Electro gryo-cator: New inertial navigation system for use in automobiles" by K. Tagami et al. (Technical Paper Series 830659, Society of Automotive Engineers, 1983), provided displayed position of the car by plotting a point on a screen. The driver could determine position by placing a transparent map over the screen. This system did not provide route directions.

The Nissan-Hitachi car navigation and information system displays position on a map, finds the shortest route to a destination taking into account real time traffic information, and gives directions by arrows on the face of a display. The system is described in "Navigation systems using gps for vehicles" by T. Itoh, Y. Okada, A. Endoh, and K. Suzuki (Technical Paper Series 861360, Society of Automotive Engineers, 1986). The system also includes a "secretary mode" which displays the driver's appointments. The system uses a CD-ROM for map data, and combines satellite positioning with dead reckoning.

"Eva: An electronic traffic pilot for motorists" by O. Pilsak (Technical Paper Series 860346, Society of Automotive Engineers, 1986) and "Digital Map Data Bases for Autonomous Vehicle Navigation Systems" by E.P. Neukirchner and W. Zechnall (*IEEE Position and Location Symposium*, pp.320-324, 1986, IEEE 86CH2365-5) describe an automobile navigation system which was developed in Germany by Blaupunkt, the University of Karlsruhe, and the federal government. The system accepts destinations as street addresses, finds minimum time routes, and gives directions by a combination of simple (arrow) graphics and voice. The system can recover from a driver error in following the route and find a new route within 50 meters of travel.

The Phillips corporation, in the Netherlands, is developing a prototype car information and navigation system called CARIN, which is-described in "Applications of the compact disc in car information and navigation systems" by M.L.G. Thoone and R.M.A.M. Breukers (Technical Paper Series 840156, Society of Automotive Engineers, 1984) and "Digital maps on compact discs" by H.J.G.M. Benning (Technical Paper Series 860125, Society of Automotive Engineers, 1986). With this system, the driver enters a destination using either a keyboard or a touch sensitive screen. The system displays routes on a map and gives spoken driving instructions. The map is stored on board in CD-ROM, and a radio link provides for updates on traffic conditions. Very little has been published about the system. It is not clear whether the CARIN system retains a map as a "vestigial" display, or because its makers do not appreciate the superiority of speech, or because they see a need for positional information other than route finding.

An article in *Automobile Electronic News* entitled "Blaupunkt adds guidance system" (Vol. 1, No. 12, Monday 3 July 1989, pg. 22.) reported that Blaupunkt (a wholly-owned subsidiary of Robert Bosch G.m.b.H.) had recently begun marketing in West Germany a CD-ROM based navigation system called Travelpilot. The system is similar to ETAK. It uses dead-reckoning and displays vehicle destination and position on a map. The map scale and position continuously adjust. There are nine possible views that can be displayed, showing distances that range from 200 meters to 50 kilometers. Currently only one CD is available with digitized maps of West Germany. It contains the street maps of 83 major cities (with populations greater than 100,000) amd all towns with populations of 500 or more. It also contains all major highways and roads with the locations of airfields, country roads, and main highways stations. Street and town names are all stored on the CD to enable the driver to enter his intended address via a simplified menu. The article noted that in the future maps will show the locations of hotels, restaurants, and castles. It reported that Blaupunkt planned to add route guidance in the future. The article makes no mention of the use of speech. It was also noted that Hertz car rental company has begun to advertise the availability of Travelpilot on some of its higher-end models and that talks are being held with some European automakers concerning OEM distribution as well.

An article in *Electronic Engineering Times* entitled "Car map system OK'd" (August 7, 1989) reported the award by the British government of a license for a vehicle guidance system. The first license was awarded provisionally to a consortium headed by the General Electric Company for a pilot scheme in central London, expected to be introduced by 1992. If the results of monitoring by the Department of Transport show that a large-scale system would not prejudice road safety or good traffic management, a second license may be negotiated which would lead to a fully commerical system covering an area of about 1000 square miles within the M25, the motorway that circles London. These commerical operations could begin by the end of 1993. It is expected that eventually, continuous guidance could be provided for most of Britain.

The system, called Autoguide, gives drivers recommended routes to their destinations using a simple graphical display fitted to the vehicle's instrument panel. The in-car unit instructs the driver where to turn and which lane to use with a simple system of arrows and other graphical symbols. The article notes that later on, other information, such as the availability of parking spaces, could be added. The article makes no mention of the use of speech. The heart of the system is the central computer which collects journey times from roads in and around the area covered. Preferred routes are continuosly updated, and this information is broadcast to subscriber vehicles from a network of strategically located roadside beacons using infrared communications techniques.

The same article notes that Autoguide is one of several approaches to road navigation in Europe. For example, the Common Market is undertaking a research program called Drive (for Dedicated Road Infrastructure for Vehicle Safety in Europe), which involves wide-ranging studies of road transport information. A similar research scheme, called Prometheus, is reportedly being operated by several European car manufacturers under the 19 nation Eureka program. Another Eureka project is Carminat, in which Dutch and French companies are developing a system that will integrate car navigation, communication, and diagnostics. The article further notes that in West Germany,

Robert Bosch G.m.b.H. and Siemans A.G. have demonstrated car navigation systems. One of the West German developments is said to be compatible with Britain's Autoguide, giving the eventual prospect of a road navigation system that could be used throughout the whole of Europe. The article notes that Britain is probably farther ahead than most countries in applying this technology.

An article in *Automotive Electronic News* entitled "U.K. Picks GEC to Head Navigation Project" (Vol. 1, No. 16, August 28, 1989, pg. 31) discusses a follow-on to Autoguide which uses roadside infra-red beacons to transmit information. Drivers reportedly enter destinations on black and white flat-screen displays, and the system provides distance, bearing, and best route. The article states that an audible signal and voice synthesizer instructs the driver when to turn or change lanes, although no details on how this is done are provided. A copy of this article is enclosed.

An article in *Automotive Electronics Journal* (Vol. 2, No. 9, April 23, 1990, pg. 22) reports that Toshiba seeks OEM partners in automaking for its navigation and guidance system. According to a Toshiba spokeman the system will have voice recognition and speech systhesis because "when you're driving, the part of the body that you're not using is the ears and the mouth." The article only discusses the intention of the manufacturer and provides no implementation details.

An article in *The New York Times Magazine* entitled "Softening of the Arteries" (August 26, 1990) discusses an experimental project called Pathfinder which involves a combination of technologies, including Travelpilot, an on-board navigational system manufactured by ETAK. Travelpilot stores map data on a disk and displays it on a screen attached to the dashboard. For the experiment, Travelpilot has been installed in 25 cars. The roadbeds of a 14-mile stretch of the Santa Monica freeway have been fitted with sensors to monitor the speed and density of traffic. This information is sent to a central computer and then relayed to the cars. There, the information is displayed symbolically on the Travelpilot computer screen. Additional data can be supplied in word messages or via digitized voice messages. In this way, the drivers are alerted to the path of least resistance. A copy of this article is enclosed.

The following discussion describes map databases, particularly with regard to features beyond those present in the DIME format.

The TIGER format, which has several improvements over the DIME format, is described in detail in the following sources: "Principal components of the census bureau's TIGER file" by J. Sobel (Research in Contemporary and Applied Geography Discussion Series 3, Department of Geography, SUNY Binghamton, 1986), "The TIGER system: Automating the Geographic Structure of the United States Census" by R.W. Marx (*Government Publications Review*, 13:181-201, 1986), "GIS, TIGER, and Other Useful Acronyms" by R.W. Marx (*National Conference of Geographic Information Systems*, Canadian Institute for Surveying and Mapping, March 1989), "Programs for assuring map quality at the bureau of the census" by R.W. Marx and A.J. Saalfeld (*Fourth Annual Research Conference*, Geography Division, Room 3203-4, Bureau of the Census, U.S. Department of Commerce, Washington DC 20233, March, 1988), "The TIGER Structure" by C. Kinnear (*AUTO CARTO 8 International Symposium on Automation in Cartography*, April 1987), and "Topology in the TIGER file" by G. Boudriault (*AUTO CARTO 8 International Symposium on Automation in Cartography*, April 1987). As

stated in the present specification, the TIGER format has several improvements over the DIME format, but is still a planar graph representing only physical connectivity. It could be used as the basis of a map database for an automobile navigation system if the extensions discussed in the specification are incorporated.

The Taxi! driving simulator, described in "Taxi! Dynamic cartographic software for training cab drivers" by M. Bosworth and R. Low (Technical report, Hunter College Department of Geology and Geography, (212)-772-4000, 1988 paper presented at the Annual Meeting of the Association of American Geographers) includes the concept of turn resistance, a number from one to ten specifying the difficulty of making a given turn, in its map database.

Neukircher and Zechnall, 1986, describes the features of the map used in the Eva system. This map has better position information than DIME. Points are stored in three dimensions and are accurate to 2.5 meters. Road segments are straight lines, chosen so that a new segment begins at either an intersection or when the change in direction exceeds 30 degrees, or when the distance from the center line exceeds 5 meters. Additional attributes of the roads include height and weight restrictions, location of magnetic anomalies, warnings, landmarks, special objects useful in descriptions (e.g. underpasses), layout of complex intersections, and signs. The map has two levels of detail. The coarse level is used for route finding, and the fine level has more detailed information for position finding. Route finding information includes two values for expected speed (one for normal conditions and a second for times of high density), the expected wait time at segment endpoints, and areas where children are likely to be playing.

The University of Calgary AVL–2000 system uses a map that originated as a Canadian government Area Master File. This format, similar to DIME, also required extensive augmentation, as described in "Digital Map Dependent Functions of Automatic Vehicle Location Systems" by C.B. Harris, L.A. Klesh, E.J. Krawkiwshy, H.S. Karimi, N.S.T. Lee (*IEEE Position and Location Symposium*, pp. 79-87, 1988, IEEE CH2675-7). Link (segment) attributes include distance, expected travel time, safety, scenic value, tolls, "impedience value" [sic], one way limitations, banned turns, road type (over– and under–pass, traffic circles, clover leaf), presence of meridians (divided?), and restricted areas. Harris describes as a "special problem" those "source of destination points which correspond to a street addresses [sic] which do not have a unique node identifier". Either their map representation cannot interpolate addresses along segments, or the route finder is restricted to finding routes to nodes only. Harris also mentions auxiliary road information including landmarks, points of interest, emergency services, commerical establishments, weather conditions, traffic flow, and road characteristics and stresses and importance of being able to update the map database over a communications link while driving.

Most systems have expanded the classification of streets. The ETAK map classification is interstate highway, semi-limited access roads and state highways, arterial, collect, light duty roads, alleys or unpaved roads, high speed ramps and low speed ramps. This rich taxonomy is essential to ETAK for choosing which roads to display (lesser roads are suppressed at larger scales to control detail) and in which colors to display them. The Eva system has a two level taxonomy: *rural*, including motorways

and federal highways with separate directional lanes and without intersections, federal highways, roads wider than six meters, roads four to six meters wide, and others, and *urban*, including divided, through, main, side, and restricted.

These maps have some questionable design decisions on the representation of legal restrictions. The ETAK map has no legal topology at all. It is not intended for route finding. The EVA map apparently encodes restrictions on turning by signs, rather than directly in the network. The Calgary map represents legal topology (one ways, banned turns) as a link attribute instead of in the network topology. It may be that the street network represents only physical topology, with the assumption that legal topology will be equivalent to the physical topology unless specially indicated.

The Back Seat Driver appears to be unique in maintaining separate but equal representations for physical and legal topology. These two topologies should be integrated because legal topology is needed for route finding, and physical topology for route description.

Some navigation systems attempt to give warnings about hazardous conditions. In those systems (Eva, Calgary), the hazards (about slope, width, or curves) are encoded explicitly into the map.

Respectfully submitted,

Sam Pasternack

Sam Pasternack
Registration No. 29,576
Attorney for Applicant

Choate, Hall & Stewart
Exchange Place
53 State Street
Boston, MA 02109
(617) 227-5020

EXHIBIT

27

FILED UNDER SEAL

EXHIBIT

28

FILED UNDER SEAL

EXHIBIT

29

FAIRCHILD PUBLICATIONS
7 EAST 12TH STREET
NEW YORK, N.Y. 10003

BULK RATE
U.S. POSTAGE PAID
FAIRCHILD PUBLICATIONS

AUTOMOTIVE ELECTRONIC NEWS

NEWSPAPER
***************** 3-DIGIT 761
AE0020827-2
ROBERT L FRENCH PRINCIPAL
R L FRENCH & ASSOCIATES
3815 LISBON ST SUITE 201
FT WORTH TX 76107

VOL. 1, NO.13 MONDAY, JULY 17, 1989 $2.50

Car Men to Put Brakes on DOD Spending

By MICHAEL G. SHELDRICK
& JIM OSTROFF

NEWS ANALYSIS

WASHINGTON — The expected nomination by President Bush of Ford executive vice-president John A. Betti as the Defense Department's undersecretary for acquisition, coming on the heels of the appointment of General Motors' vice-chairman Donald J. Atwood as deputy secretary of defense, signals a decided tightening in the operation of a government bureaucracy that has come under Congressional criticism.

Mr. Betti, who directs Ford's vast Diversified Products Operations, would fill the post vacated earlier this year by former General Motors vice-president Robert Costello, and report to Mr. Atwood, who was also president of GM Hughes Electronics Corp.

Mr. Betti is currently the subject of a standard FBI investigation that must be completed before he is offered the position. Some Detroit observers said Mr. Betti's would accept the DOD job, which reportedly had been turned down by more than a dozen other corporate executives, because it has become clear that he is not destined for a top spot when Ford vice-chairman Harold A. Poling retires in October, 1990.

Ford chairman Donald E. Petersen will retire about a year later and is likely to be replaced by Allan D. Gilmour, an executive vice-president for corporate staffs, and a member of the office of the chief executive, which includes Messrs. Petersen, Gilmour, Poling and Stanley A. Seneker, an executive vice-president and chief financial officer. Mr. Betti is part of a second tier of six other leading Ford executives.

If Mr. Betti is nominated by the President as expected, and confirmed by Congress, he would be the highest Ford executive to make the trek from Detroit to Washington since Robert McNamara was appointed Secretary of Defense in 1961.

Mr. Betti is known as one of Ford's toughest executives and most assiduous cost cutters. "We used to say 'John wears iron pants,' " said a former top Ford engineering executive who knows Mr. Betti well. "He's a very hard-nosed manager." These are all qualities essen-

See AUTO, Page 29

John A. Betti

GM, AAA FUNDING

Florida IVHS Planned

By PHIL FRAME

ORLANDO, Fla. — The U.S.'s most ambitious Intelligent Vehicle Highway System program to date — a follow-up to California's Pathfinder — is being readied for this Florida city, which is being choked with traffic from a growing population and heavy tourist travel.

The proposed $10 million, 1-year program, called TravTek, will be about four times larger than Pathfinder, a 1-year program set to begin next March in Los Angeles. In Orlando, a fleet of 100 rental cars will be equipped with the same type of enhanced communications capabilities and navigation systems to be used in Pathfinder.

The American Automobile Association, General Motors and the Federal Highway Administration (FHWA) reportedly will supply about $6 million for TravTek, according to

See FLORIDA, Page 32

Hyundai Pulls ECMs In Excel Model Recall

By MITCH BORETZ

GARDEN GROVE, Calif. — Hyundai Motor America is blaming an electronic control module for a cruise system problem that has led to a recall of approximately 8,550 Excel models.

According to a statement released by the automaker, laboratory tests found that using an optional high-end radio while the cruise control is engaged could cause a voltage spike or drain that would affect the ECM's microprocessor. As a result, the cruise control could begin a gradual acceleration or deceleration.

Fort Worth, Tex.-based Specific Cruise Systems, the OEM supplier of the cruise control ECMs, also will make a replacement control modules. Neither company would give financial details of the replacement operation, which involves about 30 minutes of labor per vehicle to install a new ECM.

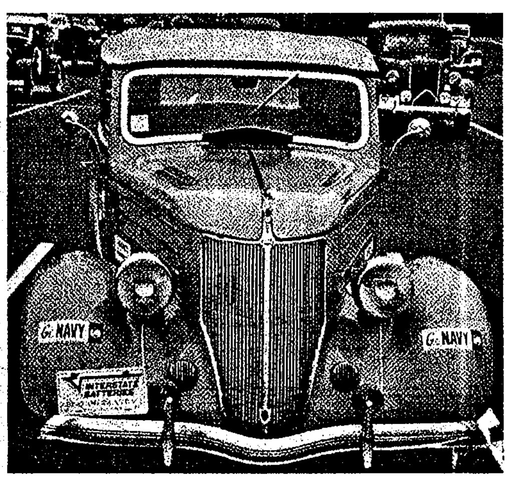
Officials at Specific Cruise Systems declined to comment on the problem except to say that the cruise controls passed tests before being delivered to Hyundai.

A spokeswoman would not say whether components other than the MPU would be affected by the problem, nor would she identify the MPU supplier. Requests for an interview with Hyundai executives were refused.

Some industry observers characterized Hyundai's recall as an example of how elec-

See HYUNDAI, Page 29

INSIDE



SANS ELECTRONICS: No ECU, ABS or active suspension, and somehow more than 100 antique cars made it across the country. More photos on Page 30. Photo by Greg Koerner

IC Makers Anxious As Car Sales Drop

By GREG GARRY

PALO ALTO, Calif. — Prospects for soft auto sales for the remainder of the year and into 1990 have sent ripples of anxiety through the semiconductor industry as IC shipments to the Big Three reportedly turned flat coming into the summer.

Reacting to a sharp decline in domestic car and truck sales — including a drop of nearly 23 per cent for the June 21-30 reporting period — General Motors, Ford and Chrysler have curtailed production with outright plant shutdowns, slowed assembly lines and lengthened model changeover periods, causing concern among semiconductor suppliers that component purchase commitments may also be cut.

While semiconductor manufacturers remain confident that dramatic increases in electronics content per car over the next 10 years will override any auto slump to provide long-term growth, a semiconductor usage report circulated by Shearson Lehman Hutton indicating flat shipments to the automotive segment in May has suppliers closely monitoring day-to-day shipments to Detroit.

As yet, no major semiconductor manufacturer is reporting a severe fiscal wound.

"We are very, very cautious right now," said Greg Williams, vice-president and director of automotive sales and marketing for Motorola's Semiconductor Products Sector in Phoenix, Ariz.

See IC, Page 32

Prototype Guidance Unit Uses Speech Synthesis

By DEAN TOMASULA

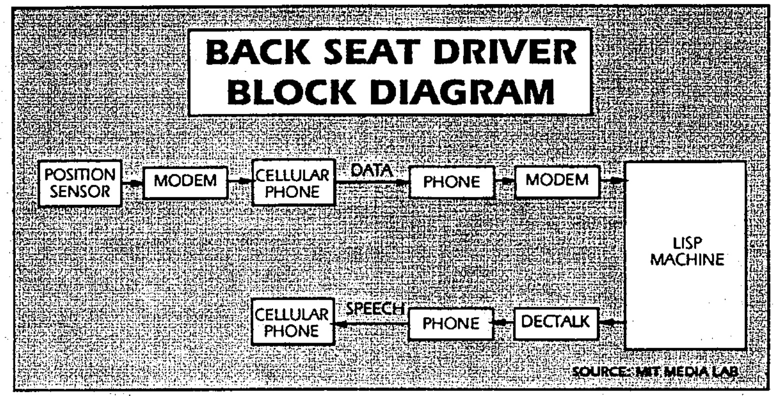
CAMBRIDGE, Mass. — Researchers at the Massachusetts Institute of Technology's Media Laboratory have developed a prototype guidance system that uses speech synthesis as a navigation aid.

The system, called the Back Seat Driver, gives directions in real time, the researchers said, in addition to planning and talking the driver through a route, warning the driver when he makes a wrong turn and advising him of an alternative route.

The Back Seat Driver, developed by MIT researchers Christopher M. Schmandt and James R. Davis, is a dead-reckoning system that uses a database stored on CD-ROM for map-matching, much like competitive navigation systems. The hardware is provided by NEC Home Electronics Ltd., which is sponsoring the project.

Two cellular telephones link a base computer to the car. The Back Seat Driver transmits the car's position and speed back to the lab's computer via modem and telephone. Since the system is only a research prototype, Mr. Schmandt pointed out, much of the computation is done by the base computer. A production system would use an on-board computer for route planning.

Speech synthesis is performed by a commercially available Dectalk text-to-speech synthesizer, which is connected to a Symbolics LISP Machine. Synthesized instructions are relayed to the driver through the second cellular phone, which is a speaker-phone. The keypad of that phone also serves as the driver's control unit for the system, allowing him to select a destination and request a repeat of previous instructions.



"MOST PROTOTYPE projects have used various forms of display to present this information, and not all have included route-finding ability," said Mr. Schmandt. "For safety reasons, a display may not be particularly suited to this task." He added that some drivers do better following spoken directions than reading a map.

He also said that when directions are being given to a driver by a passenger, the real-time aspect of the instructions becomes important. The Back Seat Driver gives "just in time" directions, taking into account vehicle speed, the difficulty of the expected maneuver, driving styles and road, weather and traffic conditions.

> "For safety reasons, a display may not be particularly suited to this task ..."

The Back Seat Driver's database currently covers 41 square miles, including parts of the Massachusetts cities of Boston, Cambridge, Brookline, Somerville and Watertown. The database is taken fromthe U.S. Geological Survey's DIME (Dual Independent Map Encoding) files.

The researchers pointed out that a DIME file alone is not sufficient for accurate route finding when used in cars. A file consists of a set of straight lines, each with a name and endpoints in latitude and longitude. Additional information such as address numbers also is included.

While the DIME files indicate physical connectivity — such as streets that intersect — it does not indicate legal connectivity. Because two streets physically intersect, it doesn't necessarily mean it is legal to drive from one to the other. For example, one of the connecting streets may run one way in the opposite direction.

THE MIT RESEARCHERS said they are currently adding landmarks to the database, such as traffic lights, stop signs, and some buildings. They also are including lane information — the number of lanes on a given road, as well as any turn restrictions on those lanes.

Since the driver cannot see the system, the Back Seat Driver periodically lets the driver know it is still working by giving warnings and reassurance. The system will acknowledge a driver's correct action with a "nice work" or "good." If the driver is approaching a turn too fast, the system will tell him to "slow down" before taking the turn.

One of the more complex set of instructions the Back Seat Driver can give is as follows: "Get in the left lane because you're going to take a left at the next set of lights. It's a complicated intersection because there are two streets on the left. You want the sharper of the two. It's also the better of them. After the turn, get into the right lane."

Mr. Schmandt said, however, that it isn't clear how much "chattiness" a driver will accept from a synthetic voice. "Certainly this feature could be useful in a rented car in a new city, where it might have some interesting things to say."

The next step in the Back Seat Driver's development, according to the researchers, is to determine exactly what a speech guidance system should say, how time and vehicle speed affect the instructions it gives, and what features a map database must have to support the generation of useful spoken instructions.□

INTRODUCTION THIS FALL

Sony to Offer Portable CD Player for In-Car Use

NEW YORK — Sony will introduce to the aftermarket this fall a portable CD player that it said was specifically designed for use in cars.

The D-180K Discman, said Shin Kobayashi, vice-president of general audio, was developed because research proved that one-third of the company's Discman units were purchased for use in cars. The D-180K also can be used as a portable CD player and at home.

The unit features a dual damping suspension to help guard against road bumps. Mr. Kobayashi said the improved anti-shock qualities of the unit make the need for a suspension mounting plate unnecessary.

THE D-180K HAS been designed for ease-of-use while driving, he said. It features oversize function controls and an orange backlit function display light and play button that are said to be easy to see in daylight or at night.

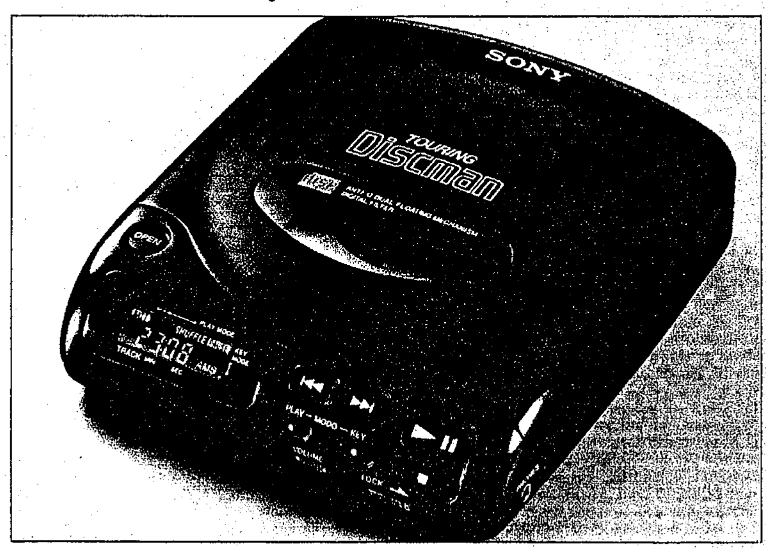
The D-180K has an oversampling digital filter to help minimize noise and improve high frequency response and features a three-spot laser optical pickup. An automatic tracking recovery mechanism disengages the laser during severe bumps and then returns it to the same spot.

THE UNIT CAN play both 5-inch CDs and 3-inch CD singles without an adaptor. It comes with a dc cable for connecting it to the car battery through the cigarette lighter. Also supplied with the D-180K are an ac adaptor for home use and a rechargeable battery.

Also supplied with the unit is a CPA-2 cassette adaptor that permits the D-180K to be played through the car's speakers. An optional CPM-70 kit allows the unit to be permanently installed in the car on a mounting arm.

The D-180K will retail for $269.95.

It will also has an optional RM-DM2 Remote Commander remote control unit for use while driving.□



GOING MOBILE: Sony's D-180K portable CD player was specifically designed for use in automobiles and features oversize control buttons.

HAR 710321

EXHIBIT

30

Telling You Where to Go

Thesis Proposal

James Raymond Davis

MIT Media Laboratory
1 January 1989

Submitted to the Media Arts and Science Section
in partial fulfillment of the degree of
Ph.D. in Media Arts and Sciences

....
Nicholas Negroponte, Thesis Advisor

....
Steve Benton, Thesis Coordinator

....
Christopher Schmandt, Reader

....
Michael Lynch, Reader
Department of Sociology, Boston University

This thesis is about helping drivers navigate through a city. This can be quite difficult. It is hard to get good instructions for getting from one place to another, and easy to get lost. A computer navigation assistant would be useful to drivers of delivery trucks, taxis, emergency vehicles; to tourists and ordinary drivers. I propose to build a computer program to give spoken driving instructions to the operator of a car in real time - that is, on a moment to moment basis. I call the system I envision the **Back Seat Driver**. This name has two connotations: first, that of an unwanted critic of one's driving skills, the second that of a helpful agent who can direct one through a locale. It is the latter sense I intend.

The next section describes earlier work on navigation systems, which I then criticise. Next I discuss in greater detail the system I intend to build, with particular emphasis on what is new and different about it. In that discussion I raise several theoretical issues which this thesis will investigate.

Previous navigation systems

Elliot and Lesk

The pioneering work on computer navigation is by Elliot and Lesk[3,4]. Finding a route through a street map is a special case of the general mathematical problem of searching a graph. Mathematicians and programmers have created algorithms for searching graphs, but these algorithms are not suitable for the practical problem of route finding. Partly this is because real street maps are large, but also simpler than general graphs. The general algorithms are expensive to use because they can handle any kind of graph. Elliot and Lesk got better performance by using simpler algorithms. More importantly, the general algorithms find the shortest route, but that is not the best route, if it is too complicated to remember or follow. The shortest route may be through a maze of shortcuts. People who saw such routes "recoiled in horror", so Elliot and Lesk modified their algorithm to prefer a route which was slightly longer but had fewer turns. This they accomplished by imposing a fictitious distance penalty of 1/4 mile for left turns and 1/8 mile for right turns. This change may even make routes faster to follow, since, at least for left turns, one may have to wait for traffic to clear before making the turn. Every known route finder includes functions to trade distance for simplicity.

Elliot and Lesk also were the first to implement a program to generate natural language driving instructions for the route. This is not a straightforward translation. The route as represented by the route-finding algorithm is a sequence of street segments, where a segment is a piece of a road chosen short enough that it is a straight line and no intersection occurs except at a segment endpoint. This

does not match any commonsense notion of a road. Route descriptions must be expressed in terms of motion along streets (across many segments) and turns. In their instructions, a route consists of a beginning, a sequence of turns and crossings (of rivers or railroads), and an ending. For each of these, there is a template to generate a sentence. The template has words fixed and others to be filled in according to the particulars. An example template is

```
Go <distance> [<intersections>] turn <direction> on <street>.
```

This template might produce `Go 0.3 miles (2 intersections) turn left on TROY HILLS RD`. Here "Go", "turn", and "on" are the fixed words, and everything else is a slot. The intersection count is optional, and only provided if relevant to the route.

A third contribution of Elliot and Lesk was to integrate the digital map with other location oriented databases, including a Yellow Pages and a personal address book. This allowed the program to find routes to addresses given a person's name, to find the closest store of a specified category, and to mention stores along the route as possible landmarks. It is not clear that this last feature was helpful, since many stores are not easy to see from the street.

Direction Assistance

My earlier project, Direction Assistance, was directly inspired by Elliot and Lesk. A long description appears in [2]. Direction Assistance differs from the work of Elliot and Lesk in several ways. It speaks its directions instead of printing or drawing a map. The interface uses only speech and touch-tone telelphone buttons, to make the program accessible from any touch tone phone, instead of requiring a computer terminal[1]. The route-finding algorithm is an A* (best first) search algorithm[7], and the route weighting scheme is different. The weighting scheme ranks roads by a four-valued "goodness" feature and penalizes routes that use less good roads by multiplying the mileage by a constant factor. It also reduces or waives the penalty for turning under a set of circumstance having to do with predicted ease of following: for example, a turn onto a one way street incurs a lesser penalty, since it is unlikely that the driver would turn the wrong way.

The third, and most significant, difference is that Direction Assistance generates high quality English prose descriptions of the route. The prose is better because Direction Assistance does not generate text directly from the route, but instead first analyzes the route into a sequence of "acts", and then describes the sequence. An act represents something that the driver does rather than motion from one street segment to the next. There are eleven different acts, each representing a different

way of moving. The type of act to use depends upon topology (how many streets are present at an intersection, and which way traffic can flow), geometry (what angles the streets make) and what kind of streets are involved. Thus we say "bear right at the fork" rather than "turn right", but we don't say that in taking an exit from a highway we are "bearing right". An act may involve more than one segment, as for instance a "U Turn" on Memorial Drive (shown in figure 1) takes one from Memorial Drive, to Danforth, and back onto Memorial Drive, yet should not be described as two successive turns. For each act there is a specialized text generator

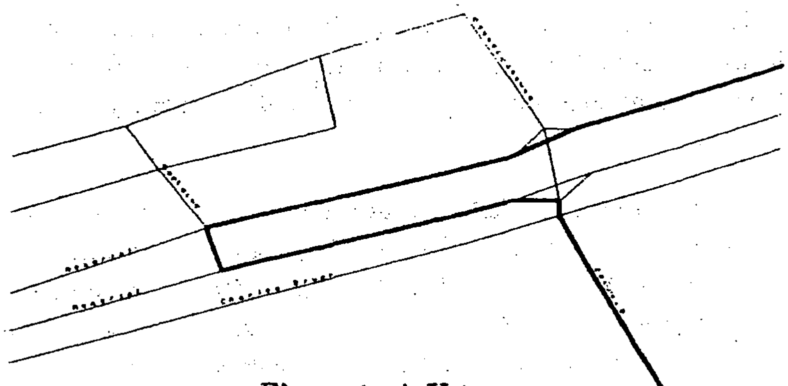


Figure 1: A U turn

to describe it, and there is a function to find an appropriate cue or landmark (e.g. a street crossed or an underpass) just before the act.

Others

Peeder Ma describes a system which gives textual directions in [11]. His work, apparently created independently, is similar to both Elliot and Lesk's and my own work. He uses A* search with a penalty factor to minimize the number of turns. Unlike Elliot and Lesk, he uses the same penalty for both left and right turns. His street map representation does not include one-way streets, or restrictions on turning ("no left turn") so it does not always find usable routes. His route descriptions use a taxonomy about as elaborate as that of Direction Assistance, but the text generated is more stylized.

The Hertz car rental company offers "Computerized Driving Directions" at some of its rental counters. The directions include approximate mileage and estimate travel time, but are highly schematic, even cryptic. An example appears in figure 2 It is not clear whether this system actually finds and describes routes independently, or simply prints out a pre-stored set of instructions.

The ETAK corporation has a navigation system for cars which displays the car's position on a map display on the dashboard. The ETAK system is complementary

```
APPROXIMATELY 16.8 MILES 0 :35 TIME
 2.0 MI NORTH  TO   I-78 WEST           enter LEFT
14.0 MI WEST   TO   NEW PRO/BERKELEY EXT bear RIGHT
                    DIAMOND HILL RD continue
 0.4 MI        TO   MOUNTAIN AVE turn RIGHT
 0.4 MI        TO   AT&T/BELL LABS on your right
```

Figure 2: Text of driving instructions provided by Hertz

to those described above: it does not tell you how to get anywhere, but does tell you where you are.

Classifying navigation systems

Navigation systems can provide two kinds of services. They can tell you where you are or they can tell you what to do. I use the adjective *positional* to refer to the first kind, and *instructional* to refer to the second. A navigation system can be both positional and instructional. Navigation systems can be further distinguished by how they provide the information to your ears (verbal) or to your eyes (visual), and by whether they work in real time or in advance. The systems of Elliot and Lesk, Ma, and Hertz provide advance, written instructions. Direction Assistance gives advance spoken instructions. The ETAK system gives real time graphic positions. An ordinary map provides neither position nor instruction information. It is not a navigation system. The categories of this classification are not independent. There can be no advance positioning system, since one can not predict the future position of the car.

There are several problems with "advance directional" navigation systems. First, they do nothing to help the driver follow the route. The driver must determine for herself when to apply each instruction. Instructions like "drive half a mile, then turn left onto Maple Street" are no use if the driver is unable to measure mileage or fails to find the street sign. Indeed, in Boston many street signs are missing. In addition, the driver must keep track of which instruction is next. A second problem is that since the instructions must be specified in advance, there is no possibility of correcting if the driver does not follow the instructions, which might happen from error, or because the instructions are wrong, or simply ill-advised (as when confronting a traffic jam).

The sole existing "positional" system, the ETAK, is graphic. One can use it for navigation, if one can read a map and find one's own route. But many people have

difficulty finding and following routes on a map[13]. Even if drivers can safely read maps while driving[5], the activity is still distracting. Drivers need to see things outside the car. A navigation aid should be verbal, and leave the eyes free for driving. Drivers are more likely to arrive at their goal when given verbal instructions then when given maps; indeed, not only are verbal instructions better than a map, but adding a map to a set of verbal instructions actually hinders performance[14].

The Back Seat Driver will provide real time spoken instructions. (It will also be capable of providing position information on request.) It will have none of the drawbacks of the other systems. It will be easier to use, and thus safer to use. The next section examines the Back Seat Driver from a variety of different perspectives, showing what work I am building on and where I expect to make a new contribution.

The planned system

The Back Seat Driver will communicate to the driver through synthetic speech. Position sensing hardware in the car will supply it with the car's position, direction, and speed several times a second. The program will find the best route to the desired destination[1], then guide the user through the route, step by step, until the user reaches the destination.

There is usually more than one way to travel between two points. The Back Seat Driver will know about people's preferences for routes. One driver may prefer the fastest route, regardless of difficulty where another may want the route that makes the least demands on attention and skills. Still another driver may prefer a scenic route than passes along the river. The route finder will be able to accommodate these preferences.

The Back Seat Driver models a route as sequence of acts taken by the driver. The task of the Back Seat Driver is to get the driver to perform each act at the right time. Driving instructions must be timed to the accuracy of about one second. (The timings would be much much tighter if the Back Seat Driver had to actually turn the steering wheel. Fortunately, a human does the actual driving.) The Back Seat Driver will give driving instructions well before the act (e.g. "You'll continue on for a mile, and then make a left onto Beacon Street.") and also just before the act ("Beacon Street is the next left, get ready.") It will attempt to anticipate the driver's mistakes (e.g. "The turn is just ahead, slow down.") and, if the driver does make a mistake, it will find a new route.

[1] I ignore the question of how the driver specifies a destination. Speech recognition would probably be best, but state of the art speech recognition does not permit this even in a quiet room, and a car is very noisy. I will use some sort of keyboard, and leave this problem to someone else.

The Back Seat Driver will have other capabilities beyond navigation. It can act as a tour guide, commenting on the streets and neighborhoods traversed. People want to be oriented while traveling - it would not be pleasant to follow a route through a completely unknown area, with only the computer as guide. Kevin Lynch writes of an *imageable* city as one where the nature of places and the relations among them are easily grasped[10]. The Back Seat Driver can help make the city more imageable by telling the driver how the paths and nodes of the city fit together.

The Back Seat Driver can also relay messages from home or office. It is already possible for people to get telephone calls while driving, but it might be more convenient to have a smart answering machine like the Phone Slave[12] screening calls. After all, we can't always safely answer the phone while driving. These messages might also include the driver's own reminders for shopping. In this case, the Back Seat Driver can keep an eye out for vendors lying along the route. It might spontaneously mention a place where needed goods could be obtained, or the driver might ask for a route to the closest provider.

Interactivity and Goals

The Back Seat Driver is a truly *interactive* system, by which I mean that both parties are active all the time. Most "interactive" computer systems are better called *reactive* - the machine and the human take turns. While the human is typing, the computer does nothing, and there is little the human can do while the machine is working, except to interrupt a calculation gone awry. Interaction is inherent in the application, because the concept of "turn taking" does not apply, since the driver's actions are continuous. There is never a time when the driver is not driving.

The Back Seat Driver pursues many goals at once. Its main goal is to get the driver to some location. This is a goal whose satisfaction requires many utterances over an extended time. Other goals include delivering messages and educating the driver about the city. These goals can not all be met at once, so it must choose which to pursue at any given moment. To make this choice properly it requires a model of action. It takes action to achieve a goal, and actions take a finite amount of time. The only kind of action Back Seat Driver can take is to talk. Since it can only say one word at a time, it must choose its words carefully. At any moment when the Back Seat Driver is not already speaking, it can either begin saying something, or it can wait. This decision must be made again and again, at each passing moment. A wrong decision can not be undone. Words once spoken can not be made unsaid, though the program can stop talking at any time; more seriously, there is nothing to be done if the program decides that it should have begun speaking two seconds previously[2]. Since the driver's actions are unpredictable, the program

[2]Save perhaps to speak faster, a possibility I will not consider further.

must improvise.

The problem, then is to allocate a resource (speaking time) to goals so that the most important goals get as much time as they need, while lesser goals get any left over. To do this, at any time, each goal must provide three pieces of information about itself:

- Is it ready to speak – that is, is there something that could be said right now which would help achieve the goal.
- If so, what is the maximum amount of time the speech would require, or
- If not, what is the minimum time until the goal will be ready to speak.

The latter two measures will be estimates, since they depend in part on the driver's future actions. Given these estimates, and a preassigned priority for each goal, the program will at each moment authorize a goal to speak, by examining each, in order of decreasing priority. The goal of highest priority that is ready to speak will be allowed to do so, unless there is some other goal of greater importance which, though not currently ready to speak, can be expected to begin speaking in less time than the the lesser goal will use in speaking. Thus the Back Seat Driver will not be constantly interrupting itself, uselessly starting a narrative it can't finish. Nevertheless, interruption will sometimes happen anyway, since the driver may make a mistake at any time, and correcting the mistake may require immediate action.

The Back Seat Driver should anticipate the driver's future actions, rather than simply react to the driver's current action. Unfortunately, the only information it has about the driver is the the position and speed of the car. The program can not tell what the driver is looking at, or see the driver's facial expression. Given limited information, it can make only limited guesses of the driver's intentions. Speed can be a useful clue. If the program knows that the driver is approaching a turn, yet is driving quickly, the program can infer that either the driver is unaware of the turn, or is unconcerned with safety. In either case, it is appropriate to tell the driver to slow down[3]. The Back Seat Driver should understand the dynamics of car motion sufficiently well that it can tell whether it is possible for the driver to slow down enough to make the turn in the time remaining. If not, the program may as well not bother, but instead begin replanning for the inevitable mistake.

[3]Comments about the driver's speed make the Back Seat Driver act more like the "nag" that the term usually connotes.

Discourse structure

The Back Seat Driver requires a model of discourse structure. A discourse is what you get when you have more than one utterance and more than one party to a conversation. This is clearly the case with Back Seat Driver, as there are two parties to the conversation, the Back Seat Driver speaking and the driver listening and reacting, and there are also multiple utterances. Discourse has structure: it has parts, and there are relations among the parts. In this section I briefly introduce discourse structure and show how it will be applied by the Back Seat Driver

I plan to use the discourse structure proposed by Grosz and Sidner[6]. They analyze discourse structure as consisting of three parts: linguistic structure, intentional structure, and attentional structure. Linguistic structure is the actual utterance, as speech or text. The elements of the discourse can be divided into a set of discourse segments which are hierarchically related – that is one segment may contain other segments. They do not specify the size of a discourse segment, but it may be as small as a phrase, though usually it is a sentence or two.

Intentional structure concerns the purposes of the discourse and its component parts. A discourse has a single discourse purpose (DP) which the speaker intends that the hearer recognize. It may have other purposes as well, but these are not treated by discourse structure. Intentional structure assigns each discourse segment a single discourse segment purpose (DSP). There may be a relation between the DSPs of any two segments. Segment A is said to *dominate* segment B when B's DSP contributes in some way to A's – that is, when B is a subgoal of A. Segment A *satisfaction-precedes* B when A's DSP must be achieved before B's.

Attentional structure concerns how objects and concepts used in the discourse become salient (brought into focus in the conversation), are used, and finally replaced. Their model of attentional structure is a stack of focus spaces, where more recently mentioned objects are higher on the stack, and thus more accessible. Not everything on the stack is accessible. When an interruption occurs, focus spaces below the stack are temporarily inaccessible.

Discourse structure is dynamic - as the conversation goes on, more and more of it is constructed. Items are pushed and popped from the attentional (focus) space, and as discourse segments are heard their DSPs are linked. At the end of a conversation, the focus space will be empty, and the intentional structure fully constructed. We can think of intentional structure as monotonically increasing - things are added, but never erased, except perhaps for corrections to mistakes. But the attentional structure is a stack - at the end of a conversation it should be empty.

Discourse structure is required for generating or comprehending fluent language. The presence of a concept in an accessible focus space is what allows one to use a

pronoun or reduced descriptor to refer to it. You might say

(1) I heard George's first speech.

(2) He's as twisted as Ronald.

Here, the object George is first mentioned in (1), and pushed on to the attentional stack. Then it can be referred to by the pronoun "he', or we could say

(3) The slippery devil was in perfect form.

Only through a mechanism like attentional structure can you tell that this sentence attributes "slippery devilhood" to George.

Intentional structure reflects the relations between concepts being related. It is as much a part of the information being transmitted as the words themselves. For instance, when a discourse contains an example, the point of it is altogether lost if the reader fails to notice the relation between the purpose of the subgoal and that of the main goal. We have a great many "cue phrases" in English which indicate intentional structure explicitly, such as "first", or "e.g.", or even "such as".) Discourse that lacks cue phrases is harder to comprehend, therefore they should be included. However, generating key phrases without a principled representation of discourse structure is quite difficult.

The Back Seat Driver must keep track of what it is doing and what it has said (or done, since it can only do by saying) so that it can repeat an utterance upon request. Sometimes the driver will not understand the Back Seat Driver's speech, since cars and streets are noisy. When asked to repeat, the system must know not merely the last sequence of words, but also the meaning and purpose of those words. Yes, sometimes a literal repeat will be enough, but in other cases the simple passing of time requires a change in wording. Consider the case when the driver is approaching a turn, and the system says "Take the second left onto Broadway.". To repeat this instruction word for word after crossing an intersection is to mislead the driver. A discourse structure is required to properly handle repeat requests.

The work proposed here will contribute to understanding of discourse structure by making principled implementation of the theories. No text generator to date has used discourse structure to motivate pronouns, reduced descriptors, or cue phrases. The Back Seat Driver will also push the limits of the theory because it operates in an environment where interruptions can occur (when the driver makes a mistake). This will ensure that the model of interruption processing described in the literature is at least workable, if not provably correct.

Text generation

A problem left unsolved by Direction Assistance is how to talk about two actions that are done in rapid succession. In ordinary speech when we make two turns in quick succession we may say "Take a left and then an immediate right." or we might call it a "jog". This latter term suggests that the two actions have been merged into a single act. Direction Assistance had a partial solution to this problem in its understanding of a U Turn (two quick turns, both in the same direction, such that one travels in the opposite direction on a same named street.) A more general solution is required.

Back Seat Driver will have to describe acts in more than one way. Direction Assistance describes every act just once, in a future, imperative tense, but Back Seat Driver will describe actions in future tense, present tense, and past tense (when describing mistakes). Descriptions of an act should depend upon context - a turn in the (far) future is just "a left" but when it gets closer it becomes "the next left". In between, it may be appropriate to identify the turn by mentioning nearby streets or landmarks, but at the time of the turn, it will probably be better to just say "now".

The text generated by the Back Seat Driver will also have several other small improvements over that of the Direction Assistant. These changes require better understanding of how to talk about roads and routes. As an example, the route Direction Assistance finds from Memorial Drive (eastbound) onto the Harvard Bridge includes a short section of Massachusetts Avenue. (See figure 1.) Even though this segment is only a few yards long, Direction Assistance dutifully includes it in the route description. What is needed here is a sense of when a street's name is significant and when it should be elided.

The Image of the City

Understanding the route

The Back Seat Driver should help drivers to understand the route it gives. This goal will make the system more pleasant to use, and will facilitate following the route, because a driver who understands the route and the city will use that knowledge to help interpret the commands Back Seat Driver gives. For example, a driver following a route that crosses the Longfellow Bridge and who knows that Main Street leads to the bridge knows about how long she'll be traveling before her next instruction, and she'll be attentive to cues at the right time, instead of having to

constantly be on the lookout.

Understanding the route has at least two aspects. A simple aspect is to be oriented with respect to the route, to know how far one has come and how much is left to go. Such information is easily provided. The Back Seat Driver knows the length of the route and the elapsed mileage, and can tell the driver on request.

A second aspect is that the route should fit into a larger model of the city. This means that the Back Seat Driver itself must have a model of the city and should speak of the route in terms that relate it to the city. There are several opportunities to do this. At the beginning of the route, the driver might hear an overview of the route, naming the major paths followed and neighborhoods crossed. During the route, locations could be described not just as street address but in larger units of neighborhoods and districts. The system might say not, "You're at 900 Mass Ave." but rather "We are now half way between Central and Harvard Squares." Orienting information can be included in instructions, or it might come between instructions, as a passing comment.

Benjamin Kuipers presents a formalization model of people's ability to learn to navigate[9,8]. In his model there are three stages of spatial knowledge:

- Sensorimotor Procedures: knowledge expressed as conditional procedures: "When you see this, do that"
- Topological Relations: containment, connection, and order
- Metrical Relations: distance, direction

Sensorimotor knowledge is the first to be acquired. Here a route is a sequence of cues and things to do. This knowledge is sufficient to follow a route, but does not support reasoning about the route. You can not reverse the route, because each cue leads only to the next, in the familiar order, and not the preceding. You may not even be able to give the route to someone else, as each cue may be recalled only in the context of the one just previous. This accounts for the familiar "I can take you there, but I can't tell you how." Back Seat Driver's instructions are sensorimotor instructions - they have to be, to be executed. We can expect that drivers can at least memorize routes at this level if they follow them several times[4].

[4] I do have a concern that people who depend completely on the system for decisions about which way to turn may never acquire this knowledge, since they will not use their own judgments. In my experience, I learn a route much faster when I'm the driver than when I'm a passenger. Partly this must be because the task of driving forces me to pay more attention to where I am, but also there must be an effect from trying to remember a previous route, or trying to interpret instructions or a map. It would be a pity if the Back Seat Driver actually impaired learning.

EXHIBIT
30

(Part 2)

People gradually acquire topological and metrical knowledge, mostly from their own experience, but they can also use other sources of knowledge. People learn metrical relations partly by seeing maps, so it is possible that they can learn topological relations by being told. This is the justification for including orienting information in Back Seat Driver.

Representing city knowledge

Back Seat Driver must represent knowledge about the city. In the terminology of Kevin Lynch [10], there are five components to a city's image: paths, districts, edges, nodes, and landmarks. A path is a channel for travel. A district is an area with some recognizable quality. A district has extent in two dimensions, and some boundaries, which may be sharp or vague. To Lynch, "recognizable quality" means that you can tell which district you are in just by looking around. An edge is a linear feature that is not useful for travel. Often, edges form the boundaries for districts. A node is either a junction of paths or a concentration of some quality. Landmarks designate a point as special in some way.

The same object may be classified in different ways by different people. The Charles River is an edge to the driver. To the boater, it is a path or even a district. As scale increases, a district becomes a node, and as it decreases, a node can be a district.

Some of these features are more essential for navigation than others. Clearly, paths are the most important feature. Knowledge about a path includes:

- Its name. A path might have more than one name. The Fitzgerald Expressway in downtown Boston is also Route 93 and Route 3 and the Central Artery. Only one of these names will be used for output, but the system should recognize any of them on input - and it should be prepared to warn the driver to expect to see or hear the other names.
- Where it goes. Paths link nodes. The nodes that a path connects are the most powerful means of placing the path. Nodes and landmarks along a path are the beginnings of the topological knowledge of the city that a driver must acquire.
- Its continuations, or, equivalently, its name changes. A single path might have different names in different places, or you might think of it as two separate paths that meet. Storrow Drive in Boston connects directly with Soldiers Field Road. This is best thought of as one path with two names. Cambridge Street in Brighton continues straight across the river into Central Square, but changes its name to River Street after it crosses the river. Is this two paths or one? It probably does not matter, as long as one makes a consistent

choice. Continuity is harder when the path forks. Lynch writes of how the continuity of Storrow Drive on the east is confused between Nashua Street and the Central Artery. Neither one is the clear successor. Sometimes the continuity can be so strong that the successor loses its identity. I always hear the Nonseignor O'Brien Highway in Cambridge called the McGrath Highway, which is its official name only in Somerville.

- Its spelling. Some names have spellings that are hard to predict from the pronunciation. If the system tells a driver to look for "Worchester" street, the driver might be looking for something like "Worster" unless otherwise informed. This problem is not limited to paths, of course, but this is the place is it going to come up the most.

Landmarks

Next in importance are landmarks, which serve as cues to tell the driver when to make a particular act. Landmarks are difficult for the program to use because it has no eyes. It is unable to tell whether a landmark is visible at any given moment. Even a large building like the Hancock building is not everywhere visible. This means that the program can never use landmarks to specify direction ("Drive towards the State House."), but only to make the current location more memorable. Mentioning the State House as one drives past it makes an "anchor" for one end of Beacon Street.

The program should only use landmarks that it can be sure are visible. The landmark need not be visible when first mentioned - it may be mentioned as a thing to expect in the route overview - but it must be visible at some time. This means that the program should use only landmarks that are immediately adjacent to the street. The program can be sure that the Charles River is visible from Storrow Drive, but not from Beacon Street. In addition, the program must distinguish day and night, since some landmarks may be hard to see at night.

Lynch also speaks of "local landmarks" – features that are unique only in the context of a portion of a route, not throughout the city. Natural spoken directions often mention features like signs, traffic lights, ordinary buildings, and other cars. The Back Seat Driver will use some of them as well. It will use street sign names implicitly, simply by naming a street. It is not always possible for a driver to read a street sign, indeed, it may be missing, but it is still good to mention the name of the street. Even if it does not help the driver to find the street, the driver wants to know what street she is on. There is no easy way to identify a path except with a name, so names are essential.

The Back Seat Driver will also use traffic lights. These come up often in ordinary

directions, and are easy to spot. Presumably the driver is already watching out for traffic lights out of desire to avoid traffic accidents and tickets, if for no other reason.

districts and nodes

There are some unsolved issues about the meaning of "district" and "node" in the Boston/Cambridge area: what to do with "squares" and how to represent "neighborhoods". Here I discuss these issues.

The Back Seat Driver will know about the "squares" of the city. There are several dozens of squares in the Boston area. (Here I'm referring to the squares people know about, not the ceremonial "squares" named by the city. In Cambrideg, at the intersection of Mount Auburn, Putnam, and Massachusetts Avenue is a tiny island of concrete which bears a sign identifying it as "Sullivan Square". But this is not the "Sullivan Square" that people know, which is in Charlestown. There must be hundreds of these squares, but I do not think any one uses them for navigation.) Very few squares are even remotely square. For the most part, each square has a unique name. But some do not. There is one "Central Square" in East Boston and another in Cambridge.

One issue is to know what constitutes a square. When can one be said to be "at" or "in" a square? In some cases, the squares have a well defined intersection. In other cases, the square extends for several blocks. A second issue is whether the term "square" denotes the intersection (a node) or the surrounding area. A "Harvard Square" address can be a half a mile from Massachusetts Avenue.

Also at issue is the meaning of the term "neighborhood". Lynch says that districts are "always identifiable from the inside" (p. 47) but many neighborhoods are not so identifiable. Unlike districts in Boston, there is nothing in the architecture to distinguish the Inman Square neighborhood from Cambridgeport. If you don't know where you are, you won't be able to tell from looking, unless you see a street sign. It may be that most of Cambridge is one "district", in which case a smaller term is needed. In some cases it might be possible to name a neighborhood for a node, as in Inman Square, but there is no defining node for Cambridgeport. It may be that the requirement for internal identification is too strong, and must be discarded.

Thesis Plan

My procedure for doing this work is empirical and iterative.

The first step is to study natural direction-giving. I will record people giving driving directions. This will be an uncontrolled study - I'll simply offer to drive people to locations they select, and record whatever instructions they give. I already have an intuition about what kinds of instructions are best to give, from my previous work with Direction Assistance. This study will add to those impressions. From this study, I'll collect a list of the forms of instructions given and the types of objects and concepts used in instructions. If I happen to make mistakes while driving, I'll study how my subjects correct me.

The second step is an iterative step. I will write the Back Seat Driver and test it. I will record the sessions, and identify places where the instructions are ambiguous, misleading, or utterly wrong, and modify the program to eliminate them. After each test, I'll learn about what works and what does not.

implementation and resources

The program will be written in Common Lisp on a Lisp Machine. The Lisp Machine will communicate with the driver through two cellular phones. One phone will carry voice from the computerto the driver, and a second will carry data back from the car to the computer. For Back Seat Driver to be practical, it would have to use a computer small enough to fit into a car. But this is not our concern at present. Using cellular phones presents some practical difficulties: the phone system makes the speech harder to understand, and data communications between the car and the computer are unreliable. These problems will make Back Seat Driver less reliable than it should be, but hopefully not so bad as to be untestable.

The program will be written in Lisp on a Symbolics Lisp Machine already owned by the Lab. There are several items required for this research – in particular, a car – but also position-sensing hardware, cellular phones, and modems. All of these will be paid for by a grant from the sponsor of this research (Nippon Electric Corporation Home Electronics Division) or will be loaned by the sponsor.

Acknowledgments

The author wishes to express his gratitude for the generous sponsorship of NEC Home Electronics, which makes this research possible.

References

[1] James R. Davis. *A voice interface to a Direction giving program.* Technical Report 2, MIT Media Laboratory Speech Group, Apr 1988. replaces the paper in the 1986 AVIOS proceedings titled Giving Directions: A voice interface to an urban navigation program.

[2] James R. Davis and Thomas F. Trobaugh. *Direction Assistance.* Technical Report 1, MIT Media Laboratory Speech Group, Dec 1987.

[3] R. J. Elliot and M. E. Lesk. *Let Your Fingers Do the Driving: Maps, Yellow Pages, and Shortest Path Algorithms.* Technical Report unpublished, Bell Laboratories, 1982.

[4] R. J. Elliot and M. E. Lesk. Route finding in street maps by computers and people. In *Proceedings of the National Conference on Artificial Intelligence*, pages 258–261, 1982.

[5] Walter W. Wierwille et al. Strategic use of visual resources by the driver while navigating with an in-car navigation display system. In *Proceedings Society of Automotive Engineers*, pages 2.661–2.675. paper number 885180.

[6] Barbara J. Grosz and Candace L. Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204, 1986.

[7] P. E. Hart, N. J. Nilsson, and B. Raphael. A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on SSC*, 4:100–107, 1968.

[8] Benjamin J. Kuipers. Modelling human knowledge of routes: partial knowledge and individual variation. In *Proceedings of the National Conference on Artificial Intelligence*, pages 216–219, 1983.

[9] Benjamin J. Kuipers. *Representing Knowledge of Large-Scale Space.* PhD thesis, MIT, July 1977. Issued as Technical Report 418.

[10] Kevin Lynch. *The Image of the City.* MIT Press, 1960.

[11] Peeder Ma. An algorithm to generate verbal instructions for vehicle navigation using a geographic database. *The East Lakes Geographer*, 22:44–60, 1987.

[12] C. Schmandt and B. Arons. A conversational telephone messaging system. *IEEE Trans. on Consumer Electr.*, CE-30(3):xxi–xxiv, 1984.

[13] Lynn A. Streeter and Diane Vitello. A profile of drivers' map reading abilities. *Human Factors*, 28:223–239, 1986.

[14] Lynn A. Streeter, Diane Vitello, and Susan A. Wonsiewicz. How to tell people where to go: comparing navigational aids. *International Journal of Man/Machine Systems*, 22(5):549–562, May 1985.

EXHIBIT

31

2007 MIT FACTS

Massachusetts Institute of Technology

FINANCIAL DATA

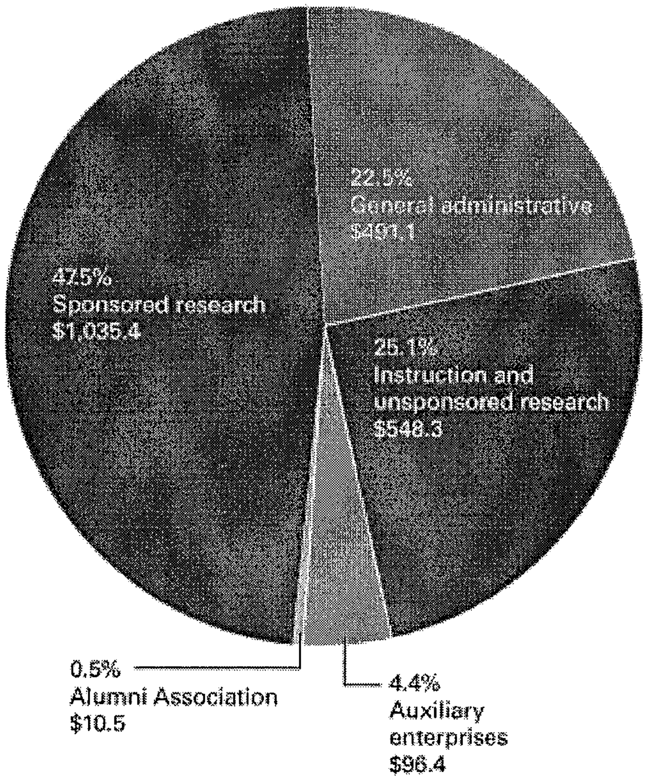
Year-end Statistics, Fiscal Year 2006 (in millions)

Value of Plant and Invested Assets	
Book value of educational plant	$1,687.8
Market value of endowed funds	$8,368.1
Book value of total investments	$7,288.7
Market value of total investments	$9,500.2

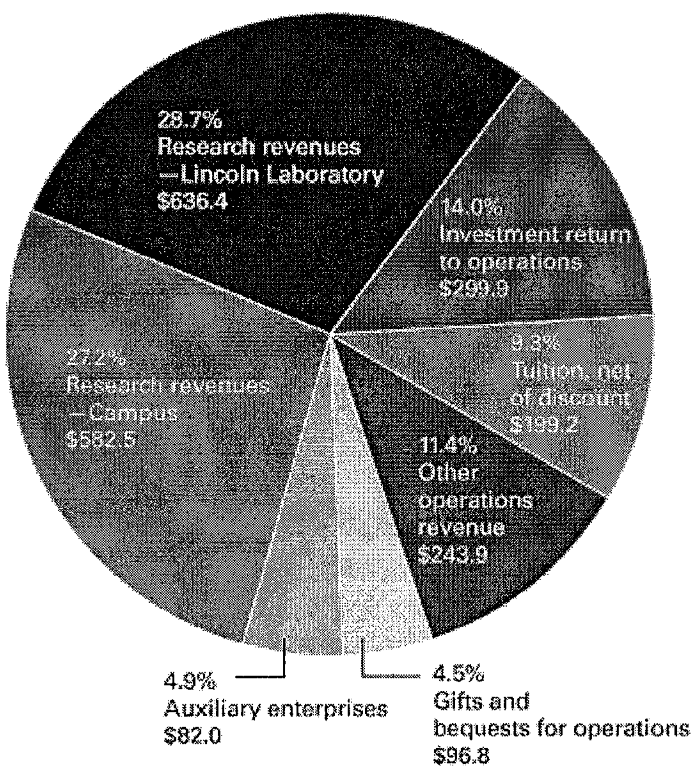
Cash Gifts to MIT	
Individuals	$114.3
Corporations	$40.2
Foundations	$86.5
Other	$0.8
Total	**$241.8**

Cash Gift Designations	
Faculty chairs	$12.2
Scholarships and other undergraduate aid	$23.0
Undergraduate education and student life	$4.6
Graduate fellowships	$20.1
Research and education programs	$133.6
Construction and renovations	$27.2
Unrestricted	$19.8
Undesignated	$1.3
Total	**$241.8**

Fiscal Year 2006
Operating Expenditures (in millions)
Total: $2,181.7 million



Fiscal Year 2006
Operating Revenues (in millions)
Total: $2,140.7 million



EXHIBIT

32

My life at the keyboard

Davis
EXHIBIT NO. 21
2/16/06
M. D. O'CONNOR

Jim Davis

Computers have been part of my culture all my life. My father worked for IBM. Every so often we'd visit him at work, and see the huge computer rooms. Sometimes he'd bring home pieces of computers for us to see. But I had no idea what these things really did. They were just cool looking. Sometime when I was 16 or 17 my father brought home some IBM documentation on programming languages and flowcharting. I tried to read them, but they did not make any sense to me - but this is not surprising, since IBM isn't exactly famous for clearly written documentation.

In 1973, I came to MIT. During the first few weeks, I went on a tour of the MIT Artificial Intelligence Laboratory, which in those days included the Logo Lab. I don't remember why I went on the tour, but I remember that I was with several other freshmen with whom I lived at my fraternity - but I can't recall if we were all together at the AI Lab because we were together as new members of the frat, or whether I came to join that frat because of liking the people I met at the AI Lab. The former seems more likely, especially since the house had a strong representation at the AI Lab. In any event, we were playing with Logo and were left pretty much alone - we had to figure it out by ourselves. We had a lot of fun with it.

So as an undergraduate I would sometimes come to the AI Lab to play (or "hack") with Logo, and I began to learn to program. I don't remember anyone teaching me - I think we must have taught each other. Sometimes I would try to use Lisp on the PDP-10 but it was too mysterious for me to figure out. That summer, I worked for IBM as an operator - a low-level position calling for about as much skill as an espresso maker. But in my spare time at work I was allowed to use the computer language APL, and this time I found a textbook for the language so I was able to learn some of it. I still had no idea how languages actually worked. I just used them.

In my sophomore year, I took the course 6.031 (which is the ancestor of the course now known as 6.001). This course explained how a computer language could be designed using a simpler language as building blocks. It also tried to give us some sense of the ideas of modularity and top down design, and most crucially, the idea of abstraction - that one can make a program which represents some concept or set of agreements, and thereafter use it without needing to know how the concept was implemented. The program becomes a "black box" whose internal details are irrelevant.

Later that year I took a second course which explained how the simplest sorts of computer languages (machine language) could be implemented by hardware circuits. I was now able to understand computer programming down to the level of individual "logic gates", if I wanted to. This reinforced my sense of the value of keeping different levels separate in order to build large, complex structure. Later, though, I would learn that one of the hardest problems is deciding where to draw the modularity lines, and that putting one's borders in the wrong place makes a system slow and difficult to use.

The next major step in learning programming was a student job at the Architecture Machine Group (which is an ancestor of the Media Laboratory). In those days, a group of students at the ArcMac were developing a new operating system for use around the lab. The operating system, being new, was full of bugs, and these in turn demanded that there be constructed many software tools for examining the structures used by the program. I had the opportunity to look over the shoulders of those who were more experienced, and even to use the tools a bit to poke around. It was while using one of these tools that I suddenly understood that there is no actual meaning in the patterns of binary ones and zeros in machine, and no significant difference between the information on a disk and in the machine's memory. A given region of memory can be an instruction, or a number, or a letter, or a picture. The difference is solely a matter of interpretation. This was perhaps the biggest "aha" in my life, and I was happy that other people were around who could understand it and why it mattered.

It was while working at this same job that I began to think not just about how to make a program do something but how to make it easy for someone else to use. This was also a step towards being a professional programmer, a worker who makes artifacts for others to use, not just for his/her own delight in making it. It was also during this time that I also began to be good enough a designer that other people started taking my ideas for design and function seriously.

After I graduated, I began to work in the real world as a programmer. My first job was at Imlac in Needham MA. The Imlac was a minicomputer (what you'd call a workstation now) that was sort of an expanded PDP-8 with a built in vector display processor. It was programmed in assembly language. Imlac's big product was a phototypsetting system, CES, which took advantage of the raster graphics to offer a kind of WYSIWYG interface for the typesetting. This was before laser printers.

I kept this first job only a year, and then moved to a new job with the (Honeywell) Multics Multics is an operating system of great historical importance. It was first developed as a partnership by MIT and General Electric as an experiment in a practical, very large time-sharing system. At the time, it was the very cutting edge of the state of the art in computer science. By the time I joined the Multics group, those days were past, but the group retained some measure of pride, and still had very high standards, even though time had passed them by. I learned several important ideas from working with the Multics people. First, my understanding of "interface" (the relation between a program and a user) expanded to include the idea that the user might be another programmer. It was important to make programs as building blocks by programmers whose needs you could not expect to easily anticipate. A second idea was that programs were meant to be read by people as well as machines. The Multics group had developed a programming process which required that all modification to the system be described and justified to a group of senior programmers before being written, and be read by some person other than the author before being installed. This was necessary because Multics was far too large for any single person to understand it. The coordination that this review board provided kept Multics stable and consistent as it grew and changed for more than 15 years. Though Multics is now nearly forgotten, it set a mark for software quality never equalled. Working with Multics taught me to be careful in my designs, to always to allow room for unanticipated future changes, and to expect people to read my programs.

In my work at Multics I came to know many people, but one of particular note was Bernie Greenberg. Bernie was one of the most brilliant programmers I have ever met. In addition, he was a very talented musician, playing both rock guitar and baroque harpsichord with equal ease, and he spoke several languages. Bernie also re-introduced me to Lisp. At that time, the MIT Artificial Intelligence Laboratory was developing the first Lisp Machines and Bernie was friends with several of the key workers on this project. I learned Lisp from Bernie not long before he left Multics to join a new startup company to commercialize the Lisp Machine. I soon left as well, to join Logo Computer Systems, a new firm which intended to implement a version of Logo for the Apple II home computer.

Logo Computer Systems was the first time I was ever with a startup firm. Instead of the formal regulations of Multics, I was with an adhoc group which included several close friends and lovers, as well as some bizarre personalities. At LCSI we worked very, very hard, because we knew that money was in short supply. We would often work for 16 to 20 hours in a row. We did almost all our work in Lisp, on Lisp Machines, and I gradually became an expert with this language. In the end, we managed to produce our product on time, but then most of us left the company as a result of political battles with the higher management.

This turned out to be a blessing though, because Alan Kay had just gone to Atari, which was then quite rich, and Alan was setting up research labs in California and Cambridge. Almost the entire Boston staff of LCSI came to form the Atari Cambridge Research Center. Atari gave us money to design the best work environment we could think of, and freedom to work on problems that interested us. Not only was I able to work on music, I was able to hire one of my friends, Tom Trobaugh, to work with me. At Atari I knew the happiness of working with a partner on problems we really cared about using the most powerful computers available. Alas, Atari began to lose money, and one day it closed the lab.

After Atari went under I enrolled in the MIT's Media Lab, as one of the first contingent of doctoral students.

The Media Lab

The Media Lab encourages students to set their own directions, in fact it insists on it. This has both pros and cons. The advantage is that you learn to be independent, to think outside the common assumptions of the field. The drawback is that you don't always have the companionship of others while learning. In my own case, I became interested in the linguistic phenomenon called "paraverbals", those inarticulate noises like "uh huh" and "hmmm" that help make conversation run smoothly. There's a pretty large literature on the subject, but I had to discover it on my own, and I'm sure it would have gone faster with a guide. On the other hand, with an experienced authority controlling my learning, I wouldn't have done what I did.

The other important thing about the Media Lab is the constant focus on demonstrating one's work. Some people complain about it, but it's very important. To do a good demo, you have to be able to explain what you're doing and why it matters to a smart but uninformed person with just a few short sentences. You have to learn to be clear, and you have to learn to express your idea for the benefit of the learner, not the

teacher. And you have to learn grace under pressure. I wish everyone learned these skills.

One of the first projects I worked on was the "phonetic dictionary". Remember when you were a kid, and you asked a grown-up for the meaning of a word, and they told you to look it up in the dictionary? That's a fine idea, except it's hard to do when you've only heard the word, and so don't know how it's spelled. This is especially true for English, with its bizarre spelling rules. The phonetic dictionary allows you to look up a word by spelling it according to how it sounds, not how it's spelled. You write some approximation of the sound of the word, and the system consults a dictionary that's organized by pronunciation. The key to the thing was being able to accept a wide range of "phonetic" spellings. For example, for the word "headache" you might write "hedayk" or "hedake". I was pleased to see this project mentioned on the very first page of Stewart Brand's book about the Lab.

My supervisor was Chris Schmandt, known to all by his login name "geek". I could write at length about his knowledge, but there are lots of smart people in the world. He has two qualities that are more rare. First, he's no autocrat. You can argue with him. There's no way to put on an air of superiority when you call yourself "geek". Second, and even more valuable, he kept his perspective. In particular, he was always taking off for a week or two at a time to bring his (then) baby daughter out into the wilderness. I learned from him that a computer will happily sit idle for a week, while a week lost from fathering is gone for ever.

My major project was the "Back Seat Driver", which was a car that could give you driving instructions in the city of Boston. It had a street map (so it knew the roads), a navigation system (so it knew where it was), and a speech synthesizer (so it could talk to you.). To actually make this work, I needed a car, and not just any car. The navigation system was supplied by our sponsor, a Japanese electronics firm, and was designed to work with only one type of car, a top of the line luxury sedan. I also needed not one but two cellular phones in the car for the communications. The Media Lab bought me what I needed and I kept the keys. I was surely the only graduate student in the USA with such lab equipment.

One day I was demonstrating the Back Seat Driver to a group from General Motors. When I took them out for a ride, they had a great time shooting pictures of each another getting into a Japanese car. Off we went driving. As usually happened, at one point the driver missed a turn. Normally, the consequence was that the BSD would calmly inform the driver of the fact and plan a new route, only in this case, the driver was a former race car driver, and he quickly made an (illegal) U Turn without even slowing down, a maneouver even seasoned Boston drivers never attempted. This caused the program to crash, but I guess that's better than crashing the car.

I'll have to add something here about the demise of Lisp Machines and the rise of Unix, and about becoming obsolete.

EXHIBIT

33

Sponsors Graphic

Search

MIT Media Lab

about us . academics . sponsors . research . publications . events . people . contact us

. sponsorship overview
. getting value from sponsorship
. intellectual property & patents
. current sponsor list

. sponsors only: insite

How Sponsors Get Value
(26 pages, updated 12-23-2004)
Download PowerPoint

Media Lab Introduction
(14 pages, updated 10-2004)
Download PDF

The Media Lab: A Unique Perspective
(1 page, updated 02-2005)
Download PDF

How to Get Value from Media Lab Sponsorship

There are many ways to leverage the Media Lab's resources:

Visit, view, and discuss. At our twice-yearly consortium meetings, hundre of working prototypes developed at the Lab demonstrate how people and technology will work together in the future. At these events, you are welco to suggest specific research projects or new directions for ongoing projects. In addition to attending consortium meetings, you are encouraged to:

- **consult** faculty and research assistants (both doctoral and master's students) from more than 30 research groups;
- **attend** special symposia and workshops;
- **visit** the Lab during the year for individual discussions and demonstrations relevant to your company's specific research agenda;
- **host** faculty and research assistants at your company, to learn more about what you do and to provide you with updates on the latest technologies and their relevance to your company.

Keep informed with immediate access to the Lab's latest developments through **insite**, our password-protected hub for internal sponsor communications (http://www.media.mit.edu/insite). **insite** consolidates technical notes on research projects, live Webcasts of Lab events, an archiv of Media Lab colloquia and meetings, and recent theses. FRAMES, our bimonthly newsletter, also provides Lab news, as do frequent email pointer to stories of specific interest to the Media Lab community.

Make use of our intellectual property. Get a two-year jump on non-sponsors for licenses to Lab technology—at no additional cost. You can als take code and prototypes into your own shop and adapt them to your needs.

Initiate peer contact. An important part of belonging to one of the Lab's consortia is the opportunity to share information in a neutral environment, and to form alliances (a frequent occurrence). As a consortium member, yo are exposed to a variety of companies and individuals from the US and abroad.

Evaluate. Media Lab research shows what works and what doesn't; what's hot and what's not; what's going on in industry and what you can expect—through communication with our community of sponsors around the countr and the world, and through connections to research colleagues and sponsor

Stay steps ahead by using the Lab as a window to investments and start-u

More than 60 start-ups have been created by former students since the Lab began. Many projects at the Lab have evolved into products. Sponsors have an inside track on potential opportunities.

Scout new talent. Perhaps the most significant resource is the Lab's potenti labor pool: students. With additional funding, our graduate fellows progran gives sponsors the opportunity to connect with specific students and researc groups in areas of particular interest to the company. Student fellows can carry your company name and can rotate annually. But any Lab student can be hired as a summer intern or upon graduation. You know who they are an what they can do. You know they're smart.

EXHIBIT

34

FILED UNDER SEAL

EXHIBIT

35

FILED UNDER SEAL

EXHIBIT

36

FILED UNDER SEAL

EXHIBIT

37

THAM 9.3

CD-ROM ASSISTED NAVIGATION SYSTEM

* Osamu Ono
Hideni Ooe
Masahiro Sakamoto

* 2nd R & D Department, R & D Laboratoriese, NEC Home Electronics, Ltd.
262, Hisamoto Takatsu-ku, Kawasaki, Kanagawa, 213 Japan.

1. ABSTRACT

A path-induction navigation system with aural direction has been developed by using CD ROM and aural information for versatile and safer navigation. CD ROM ensures high-precision self-sustained navigation and provide multiple mobile information.

2. INTRODUCTION

In the several years since its first implementation, the mobile navigation system has been the core of mobile information system. The heart of system is a database containing information that does not change, stored on CD ROM. This article introduce a high-precision navigation system that uses CD ROM and path induction navigation system that uses aural information for excellent accuracy.

3. NAVIGATION SYSTEM

Modern navigation systems largely fall into two types:self-contained navigation systems capable of locating a car position without external assistance; and radio navigation system that locate a car with external assistance (i.e. radio waves).

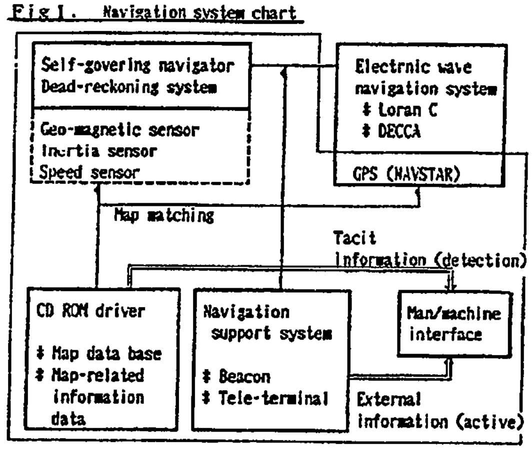
Self-contained systems are said to use inertial navigation or dead-reckoning. They acquire data through a magnetic sensor (for the Earth's magnetic field), a wheel sensor, or an internal gyro-sensor and processes sensor data to determine the car's position.

Radio navigation systems include beacon systems, loran C, GPS, etc. In particular, GPS range from a one-channel time-division reception system (capable of locating an absolute position, by receiving radio wave from four satellites to determine the three-dimentional position, velocity, etc.) to a five-channel simultaneous reception system including a spare channel.

Self-contained systems have a problem in that the absolute position cannot be determined and errors may be cumulative. Radio navigation systems have drawbacks too; errors or no reception due to the receiving environment, and rather high cost.

We have developed a new system that uses a high-precision sensor. The system implements an inquiry system for map data (map-matching) through CD ROM to reduce cumulative error. This system is a hybrid that will allow future extension, for example a beacon receiver to determine absolute position, derive external information, and use GPS input.

Fig.1. Navigation system chart



4. CD ROM

In locating a car with high accuracy, it is important to reduce sensor error. But cumulative error is inevitable over a long drive. Self contained systems require a map-matching function to prevent such error, but this requires accurate map data. To provide map data, we have used CD ROM with a capacity of 540M bytes, best suited to store a navigation database.

Further the unit can be quite effective for on-board entertainment if used in conjunction with a mobile CD. Information on CD ROM for a navigation system can be classified in two categories;

CH2564-3/88/0000-0118 $1.00 © 1988 IEEE

(1) Map data
(2) Service data

Map data is essential to the navigation function, while service data are linked to various points on the map to provide a variety of convenience for the user.

(1) Map data

Map data consists of picture data and map-matching data. Picture data is used to draw a map on the display, and allows several layers of information. For example, the background layer is for drawing in rivers, parks, building, etc. Other layers include a character layer for names of places and streets on the map, and a road layers for roads.

Map-matching data is used to describe road crossings (link information) crossings and position of curve (node information) with the node (i.e., to describe the road network). These data are used to determine where your car is.

(2) Service data

Maps contain a variety of information in addition to topography and the names of places, depending on the particular applications. Like a town map, the information also includes various data related to positions. For example,

* Descriptions, fees, etc. of interchanges and service areas along expressways.
* Positions, fees, sizes of parks
* Information on traffic restrictions, such as one-way, no-entry, and no right(left) turns.
* Detours in case of traffic jams.
* Location and description of gas stations, hospital, hotels, etc.
* Sightseeing hints.

This information can be provide both visually and by audio.

5. Path induction system through voice

This system was implemented with integrated map data and service data. We used map-matching of the map data, voice data, and position data. Data are entered interactive by using the display. The user enters the destination and crossings through which you want to go and where you want to turn. Just turn on the system and drive. The system will explain your the optimum route to you by voice. The system provide you safer navigation because you need not watch the display.

The block diagram of this system and the data structure diagram of CD ROM are shown below.

Fy2 OUTLINE OF NAVIGATION SYSTEM

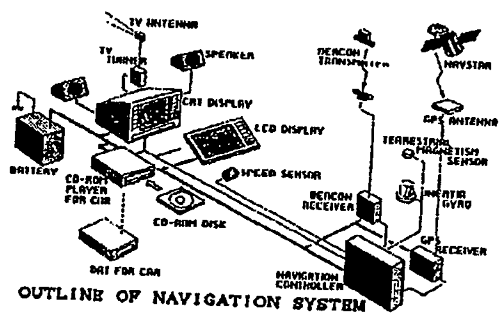
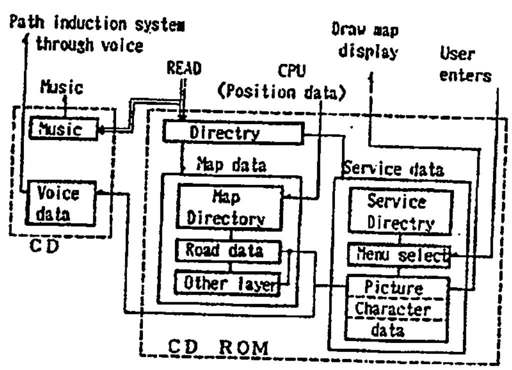


Fig 3. The data structure diagram of CD ROM



HAR 093767

EXHIBIT

38

FILED UNDER SEAL

EXHIBIT

39

FILED UNDER SEAL

EXHIBIT
40

FILED UNDER SEAL